Factors affecting the distribution of three non-indigenous riparian weeds in north-east England

Willis, Stephen Geoffrey

Factors Affecting the Distribution of Three Non-indigenous Riparian Weeds in North-East England

by

Stephen Geoffrey Willis B.Sc. M.Sc.

Department of Biological Sciences,

University of Durham.

1999.

This thesis is submitted in

candidature for the degree of

Doctor of Philosophy

Declaration

The material contained within this thesis has not previously been submitted for a degree at the University of Durham or any other university. The research reported within this thesis has been conducted by the author unless indicated otherwise.

Abstract

The work presented uses a multi-disciplinary approach to examine the factors important in determining the distribution of the non-indigenous species, *Impatiens glandulifera, Heracleum mantegazzianum* and *Fallopia japonica* at a river catchment level.

Distribution data for all three species along the Wear catchment, Co. Durham, were initially collected and the distribution of the species, in terms of density and abundance in different zones and habitats of the riparian system, were investigated. This work concluded that zones of the riverbank were used to differing extents by the three species. For all three species the lower riparian zone was the most important for the occurrence of populations.

Data extracted from the Environment Agency's River Corridor Survey were used to provide information on characteristics of two river catchments. Examination of these data in association with the alien species distribution data highlighted differences in distribution patterns related to factors such as woodland, ruderal vegetation and bank management.

Modelling species occurrences using the RCS data produced good predictive models for the two seed producing species (*Impatiens* and *Heracleum*) within a catchment but only poor models for *Fallopia*, with its solely vegetative method of spread. However testing such models on alternative catchments resulted in a reduction in predictive ability; the best overall models being derived from data amalgamated from both catchments. Variables selected in the models were found to concord with habitat preferences given elsewhere and also highlighted the importance of climate. Increasing the resolution of the collected data from 500m to 50m sections was found not to greatly improve the ability to predict species presence, though these data did allow predictions of *Impatiens* abundance to be made.

Demographic analyses in different habitat types emphasised the importance of herb/ruderal vegetation, though all three species were found to persist in woodland areas despite reduced productivity. Other experiments examining the effects of climate, as represented by altitude, on the performance of the study species indicated that factors such as seed production and plant biomass varied with altitude, whereas germination did not.

The thesis highlights potential shortfalls in producing predictive models for non-indigenous species based on non-equilibrium distributions and demonstrates interesting scale-dependent phenomena. It is suggested that whilst *Impatiens* may be largely climatically limited, *Heracleum* and *Fallopia* are more likely to be dispersal limited.

Acknowledgements

Many thanks must be offered to my supervisor, Phil Hulme, without whose constant improvements, revisions and amendments I would have been able to leave Durham a more sane, but also much more incompetent biologist. My deepest thanks must also be offered to my second supervisor, Brian Huntley, for his balanced opinions, his unerring ability to highlight problems and his willingness to drop everything to solve even trivial problems.

My thanks also go to the other members of the research consortium who have assisted in various ways and offered ideas and suggestions. Richard Wadsworth and Yvonne Collingham were very helpful sources of computing assistance. Thanks also to Yvonne for helping in the arduous task of recording species distributions. Jane Croft and Chris Preston deserve many thanks for manning an additional experimental plot for me down at Monks Wood. Unfortunately this data couldn't be included in the climate chapter here. I can sympathize with Jane over a summer of *Impatiens* pod-bagging.

Judy Allen of the Environment Change Research Group deserves many thanks for assisting in all problems great and small. Similarly Andy Joyce's wizardry at interpolating climates saved me months of background reading of climatic modelling.

The Environment Agency were a great help in providing River Corridor Survey data, in particular Colin Blundell and Katy Dickson. Similarly Northumbrian Water and English Nature deserve thanks for allowing me to set up experiments on their sites.

I wouldn't have been able to write up this thesis without the help and tolerance of Louise. Thanks for putting up with the long hours and also the provision of pizza when it was needed. The lab 15 caffine crew also deserve thanks, though more for their water boiling skills than their academic assistance. My full figure owes much to Liam and Yih-Yia.

My brother unwittingly helped by donating his PC to the cause, though had he known how long it would remain resident at Durham I'm not so sure he would have offered its services.

Finally my thanks must go to my parents, to whom this thesis is dedicated, without whose constant assistance and financial support over the years I would never have made it this far. I am particularly indebted to my father for encouraging me in all things ornithological, piscatorial and botanical over the years. All those mornings fishing on the Tees and visiting Teesmouth and Teesdale finally paid off.

Table of Contents

Declaration ii
Abstract iii
Acknowledgements iv
List of Figures x
List of Tables xiii

CHAPTER ONE – Aims and Background

1.1 Invasions and the study of non-indigenous species 1
1.1.1 What is a non-indigenous species? 1
1.1.2 An introduction to the problems associated with alien invasions. 1
1.1.3 Background to the study of alien plant species. 2
1.1.4 Approaches to the study of alien plant invasions. 3
1.2 Aliens in the British flora. 3
1.2.1 Why are river systems so invasible?. 4
1.3 The study of riparian invaders – The study species 5
1.3.1 Brief history and life history of the three species. 5
1.3.2 *Impatiens*. 5
1.3.3 *Heracleum*. 6
1.3.4 *Fallopia*. 8
1.3.5 The invasive problems of the study species with particular reference to their spread in the British Isles. 8
1.4 Scope and aims of the thesis 9
1.5 Plan of the thesis 10

CHAPTER TWO – Introduction to the study area

2.1 Climate, geology and habitats 12
2.2 Durham's alien flora 15
2.3 The County Durham rivers 16

2.4 Vegetation and ecology of the riparian environment 17
2.5 Riverine species & plant communities in County Durham 17

CHAPTER THREE - The River Corridor Surveys

3.1 Aims and Introduction 22
3.1.1 What are the River Corridor Surveys? 22
3.1.2 RCS data collection methodology 22
3.1.3 The Tees and Wear catchment RCS 23
3.2 Methods 27
3.2.1 Data extraction methods for the RCS data 27
Details of the database 27
Bankside management 28
Land-use, substrates and riverbank features 28
Macrophyte data 29
3.2.2 Data reduction 29
3.2.3 Analysis techniques 30
3.3 Results 32
3.3.1 Catchment characterisations derived from RCS variables 32
3.3.2 Species distributions 35
3.3.3 Summary statistic 40
3.3.4 Management regime 43
3.3.5 Morphological features 44
3.3.6 Bankside features 45
3.3.7 Surrounding land-use 46
3.3.8 Composite variables 48
3.4 Discussion 49

CHAPTER FOUR - Modelling using the River Corridor Survey Data

4.1 Aims and Introduction 53
Riparian species modelling 53
4.2 Materials and Methods 54
4.2.1 Autocorrelation: testing and elimination 55

Associations in the RCS data 55
Approaches to eliminate autocorrelation 56
a)The study species 56
b)The RCS variables 57
4.2.2 Logistic regression methodology 58
4.2.3 RCS macrophyte data 60
4.3 Results 61
4.3.1 Modelling *Impatiens* on the Wear and Tees 61
Modelling within the Wear catchment 61
Modelling within the Tees catchment 61
Testing *Impatiens* models within catchments 64
Testing *Impatiens* models between catchments 64
Amalgamated *Impatiens* models 65
4.3.2 Modelling *Heracleum* on the Wear and Tees 67
Modelling within the Wear catchment 67
Modelling within the Tees catchment 68
Testing *Heracleum* models within catchments 70
Testing *Heracleum* models between catchments 70
Amalgamated *Heracleum* models 71
4.3.3 Modelling *Fallopia* on the Wear and Tees 73
Modelling on the Wear catchment 73
Testing models within and between catchments 75
Amalgamated *Fallopia* models 76
4.3.4 Modelling species distributions using subsets of the RCS database 77
4.4 Discussion 78

CHAPTER FIVE – Fine-Scale Catchment Modelling
5.1 Aims and Introduction 86
Refining the RCS data 86
5.2 Methods 89
5.2.1 Defining the recording units for fine scale modelling 89
5.2.2 Timing and logistics of the survey 92
5.2.3 Descriptive statistics 92

5.2.4 Modelling using the refined distribution data and the RCS variables 93
5.2.5 Modelling with the 50m distribution data and the 50m habitat data 93
5.3 Results 94
5.3.1 Bank zone recording 94
5.3.2 Species occurrence in the three bank zones 94
5.3.3 Abundance distribution 99
5.3.4 Density distribution 100
5.3.5 Using the refined distribution data and the RCS variables 102
5.3.6 Modelling using species abundance 106
5.3.7 Analyses using the 50m resolution data of both species and habitat type 108
5.4 Discussion 111

CHAPTER SIX - Species Performance in Riparian Habitats
6.1 Aims and Introduction 119
6.2 Methods 120
6.2.1 Experimental design 120
6.2.2 Experimental protocol 120
6.2.3 Analysis 121
6.3 Results 123
6.3.1 Seedling numbers 123
6.3.2 Disturbance treatment 125
6.3.3 Growth season mortality 127
6.3.4 Is mortality density dependent? 131
6.3.5 The effects of density-dependent mortality over the growing season 132
6.3.6 Over-winter mortality 133
6.3.7 Reproduction 134
i) *Impatiens* 134
ii) *Heracleum* 136
6.4 Discussion 139

CHAPTER SEVEN - Species Traits along a Climatic Gradient
7.1 Aims 146
7.2 Methods 146

7.2.1 Site locations and layout 146
7.2.2 Experimental design 147
7.2.3 Climate variables 150
i) Interpolated climates 150
ii) Meteorological data 152
7.3 Results 154
7.3.1. Germination 154
a) Timing of germination 154
b) Percentage germination 158
7.3.2. Growth 160
7.3.3. Survivorship 163
7.3.4. Biomass 165
7.3.5 Flowering & seed production 169
a) Timing of seed set 169
b) Seed output 171
c)*Impatiens* seed weight 175
7.3.6 Herbivory 176
7.4 Discussion 177

CHAPTER EIGHT – General Discussion
8.1 The role of habitat type 188
8.2 The importance of climate 190
8.3 What determines distributions? 191
8.4 Comparing models with demographic analyses 192
8.5 The role of scale in species modelling 193
8.6 The results in light of published work 193
8.7 Further work 196

References 198
Appendix I 222
Appendix II 228

List of Figures

Figure 1.1 Distribution of *Fallopia*, *Heracleum* and *Impatiens* in the British Isles.....7
Figure 2.1 Map showing the location of County Durham and its major rivers.....13
Figure 2.2 Typical cross-sections of the lowland reaches of the Durham rivers 20
Figure 2.3 Typical cross-sections of the upland reaches of the Durham rivers 21
Figure 3.1 Map of County Durham showing the major drainage channels. 24
Figure 3.2 Example of a reasonable quality RCS section diagram. 25
Figure 3.3 Example of a poor quality RCS section diagram. 26
Figure 3.4 The bankside characteristics of the Tees and Wear catchments. 33
Figure 3.5 The land-use characteristics of the Tees and Wear catchments. 34
Figure 3.6 The riverine substrates of the Tees and Wear catchments. 35
Figure 3.7 Distribution of *Impatiens* along the Wear catchment 37
Figure 3.8 Distribution of *Impatiens* along the Tees catchment 37
Figure 3.9 Distribution of *Heracleum* along the Wear catchment 38
Figure 3.10 Distribution of *Heracleum* along the Tees catchment 38
Figure 3.11Distribution of *Fallopia* along the Wear catchment 39
Figure 3.12Distribution of *Fallopia* along the Tees catchment 39
Figure 3.13 *Heracleum* on the Wear in relation to morphological features.....44
Figure 2.14 *Impatiens* on the Tees in relation to morphological features. 45
Figure 2.15 *Heracleum* on the Wear and Tees in relation to bankside features. 46
Figure 2.16 *Impatiens* on the Tees catchment in relation to bankside features. 47
Figure 2.17 *Heracleum* on the Wear and Tees in relation to surrounding land-use.....47
Figure 2.18 *Impatiens* on the Tees catchment in relation to surrounding land-use.....48
Figure 2.19 *Fallopia* on the Tees catchment in relation to surrounding land-use. 48
Figure 2.20 *Heracleum* on the Tees catchment in relation to composite variables. 48
Figure 2.21 *Impatiens* on the Tees catchment in relation to composite variables. 49
Figure 4.1 χ^2-tests of the occurrence of the species in an RCS section in relation to upstream occurrence.....56
Figure 4.2 Predicted distributions of *Impatiens* along the Tees, using the best Wear model 65
Figure 4.3 Predicted distributions of *Impatiens* along the Tees, using the amalgamated Tees/Wear model 66
Figure 4.4 Predicted distributions of *Heracleum* along the Tees, using the best Wear model. 71

Figure 4.5 Predicted distributions of *Heracleum* along the Tees, using the amalgamated Tees/Wear model. 72
Figure 4.6 Distribution of the road/bridge variable along the Wear catchment. 82
Figure 4.7 Occurrence of a) disturbed areas and b) *Heracleum* along the Wear. 83
Figure 5.1 Cross-section showing the three zones of the riverbank 89
Figure 5.2 Absence of the upper- and non-riparian zone along the Wear 94
Figure 5.3 Proportion of 50m cells occupied by *Impatiens* in the 500m RCS sections .95
Figure 5.4 Proportion of 50m cells occupied by *Heracleum* in the 500m RCS sections.96
Figure 5.5 Proportion of 50m cells occupied by *Fallopia* in the 500m RCS sections. ...97
Figure 5.6 Occurrence of the study species in the three bank-height zones 98
Figure 5.7 50m sections occupied within 500m sections containing a study species. 98
Figure 5.8 Abundance of *Impatiens* in the three bank-height zones. 99
Figure 5.9 Abundance of *Heracleum* in the three bank-height zones. 99
Figure 5.10 Abundance of *Fallopia* in the three bank-height zones. 100
Figure 5.11 Recorded densities of *Impatiens* in the three bank height zones. 101
Figure 5.12 Recorded densities of *Heracleum* in the three bank height zones. 101
Figure 5.13 Recorded densities of *Fallopia* in the three bank height zones. 101
Figure 5.14 Abundance of *Impatiens* against predictions using RCS variables 107
Figure 6.1 Germination of *Heracleum* seedlings in the habitat quadrats. 124
Figure 6.2 Germination of *Impatiens* seedlings in the habitat quadrats. 124
Figure 6.3 Maximum number of *Fallopia* shoots in the habitat quadrats. 124
Figure 6.4 Germination and survival of *Heracleum* seedlings in disturbed and undisturbed vegetation. 125
Figure 6.5 Germination and survival of *Impatiens* seedlings in disturbed and undisturbed vegetation. 126
Figure 6.6 Growth and survival of *Fallopia* shoots growing in disturbed and undisturbed vegetation. 127
Figure 6.7 Survivorship of *Impatiens* seedlings grown in different habitats 128
Figure 6.8 Mean survival time of *Impatiens* individuals in different habitats. 129
Figure 6.9 Survival of *Heracleum* seedlings in different habitats 130
Figure 6.10 Mean survival time of *Heracleum* individuals in different habitats 130
Figure 6.11 Survival of *Heracleum* seedlings in woodland (1998) at time t+1 against density at time t. 131
Figure 6.12 Survivorship of *Impatiens* seedlings in ruderal habitat (1998) at time t+1 against density at time t. 132
Figure 6.13 Survival of *Impatiens* seedlings in ruderal vegetation (July 1998) 133

Figure 6.14 Summer and winter mortality rates for *Heracleum*. 134
Figure 6.15 Mean seed output of *Impatiens* plants in different habitats. 134
Figure 6.16 Mean density of flowering *Impatiens* plants/m^2 in different habitats 135
Figure 6.17 Mean numbers of seeds in *Impatiens* pods from different habitat types. .. 135
Figure 6.18 *Impatiens* seed estimates per square metre 136
Figure 6.19 Mean rays on terminal umbels of *Heracleum* plants in different habitats. 137
Figure 6.20 Mean number of satellite/axillary umbels on *Heracleum* plants in different habitats 137
Figure 6.21 Seeds on primary umbels of *Heracleum* plants in different habitats. 137
Figure 6.22 Seeds on satellite umbels of *Heracleum* plants in different habitats 138
Figure 6.23 Total seed output of *Heracleum* plants in different habitats 138
Figure 7.1 Locations of experimental sites along the Wear catchment. 147
Figure 7.2 Climatic data from the TIGER dataset for the study sites. 151
Figure 7.3 Derived climatic parameters for the transect sites over the study period. ... 153
Figure 7.4 Germination dates of *Impatiens, Heracleum* and *Fallopia* propagules 155
Figure 7.5 Germination time of *Impatiens, Heracleum*, and *Fallopia* propagules in relation to spring frost days. 156
Figure 7.6 Germination time of *Impatiens, Heracleum* and *Fallopia* propagules in relation to Spring heat sum 157
Figure 7.7 Maximum germination of *Impatiens* and *Heracleum* seedlings and *Fallopia* shoots across the experimental transect. 158
Figure 7.8 Proportion of *Impatiens* seeds surviving to emergence against winter frost days. 159
Figure 7.9 Maximum heights of *Impatiens* plants at the study sites. 160
Figure 7.10 Maximum heights of *Fallopia* plants at the study sites. 161
Figure 7.11 Maximum heights of *Heracleum* plants at the study sites. 161
Figure 7.12 Maximum heights of the species in relation to summer heat sum 162
Figure 7.13 Survival of *Heracleum* seeds across the experimental transect. 164
Figure 7.14 Survival of *Impatiens* plants along the experimental transect. 165
Figure 7.15 Biomass of 1st-year *Heracleum* plants grown at the study sites. 166
Figure 7.16 Biomass of 2 year-old *Heracleum* plants grown at the study sites. 166
Figure 7.17 Biomass increase of *Fallopia* plants grown at the study sites. 167
Figure 7.18 Weight gain of 1st year *Heracleum* plants against summer heat-sum 168
Figure 7.19 Biomass of two year-old *Heracleum* plants against summer heat-sum. 169
Figure 7.20 Weight gain of *Fallopia* plants in relation to summer heat-sum. 169
Figure 7.21 Timing of flowering and pod production by *Impatiens* at the study sites .170

Figure 7.22 Mean pod production of *Impatiens* plants across the altitudinal gradient. 172
Figure 7.23 Pod production of *Impatiens* plants in relation to summer heat-sum 172
Figure 7.24 Temporal changes in seed content of *Impatiens* pods from plants across the altitudinal gradient.....174
Figure 7.25 Temporal change in seed content of *Impatiens* pods from inflorescences.174
Figure 7.26 *Impatiens* seed weights over time across the altitudinal gradient.....175
Figure 7.27 *Impatiens* seed weight in relation to preceding monthly heat-sum 176
Figure 7.28 Growth season heat-sums of altitudinal sites compared to predicted climatic limits for *Fallopia* and *Impatiens*.....185

List of Tables

Table 2.1 Sunshine hours during haar conditions and associated monthly means.....15
Table 3.1 Features recorded in RCS maps. 23
Table 3.2 Basic variables recorded for sections of the River Corridor Survey 27
Table3.3 Composite variables created to amalgamate related groups of variables.....30
Table 3.4 RCS variables that are significant/insignificant in relation to the study species' occurrences. 40
Table 4.1 The occurrence of study species on opposite banks of the river 55
Table 4.2 Interpretation of Cohen's κ statistic. 58
Table 4.3 Logistic regression models using macrophytes as predictor variables.....60
Table 4.4 Logistic models produced for *Impatiens* using the Wear1 dataset.....61
Table 4.5 Logistic models produced for *Impatiens* using the Wear2 dataset.....61
Table 4.6 Logistic models produced for *Impatiens* using the Tees1 dataset 61
Table 4.7 Logistic models produced for *Impatiens* using the Tees1 dataset.....62
Table 4.8 Variables used in the various *Impatiens* logistic regression equations. 63
Table 4.9 Test of the most parsimonious *Impatiens* model to predict occurrence in the remaining sections of a catchment.....64
Table 4.10 Test of the most parsimonious *Impatiens* model to predict occurrence in the alternate catchment. 64
Table 4.11 The ability of *Impatiens* models build using the Wear & Tees data to predict occurrence in the remaining sections of the two catchments.....66
Table 4.12 Variables in models built to predict the occurrence of *Impatiens*.....66
Table 4.13 Logistic models produced for *Heracleum* using the Wear1 dataset.....67
Table 4.14 Logistic models produced for *Heracleum* using the Wear2 dataset.....67
Table 4.15 Logistic models produced for *Heracleum* using the Tees1 dataset 68
Table 4.16 Logistic models produced for *Heracleum* using the Tees1 dataset.....68

Table 4.17 Variables used in the various *Heracleum* logistic equations. 69
Table 4.18 Test of the most parsimonious *Heracleum* model to predict occurrence in the remaining sections of a catchment. 70
Table 4.19 Test of the most parsimonious *Heracleum* model to predict occurrence in the alternate catchment. 71
Table 4.20 The ability of *Heracleum* models build using the Wear & Tees data to predict occurrence in the remaining sections of the two catchments. 72
Table 4.21 Variables in models built to predict the occurrence of *Heracleum*. 73
Table 4.22 Logistic models produced for *Fallopia* using the Wear1 dataset. 73
Table 4.23 Logistic models produced for *Fallopia* using the Wear2 dataset. 74
Table 4.24 Variables used in the various *Fallopia* logistic regression equations. 75
Table 4.25 Test of the most parsimonious *Fallopia* model to predict occurrence in the remaining sections of a catchment. 75
Table 4.26 The ability of *Fallopia* models build using the Wear & Tees data to predict occurrence in the remaining sections of the two catchments. 76
Table 4.27 Variables in models built to predict the occurrence of *Fallopia*. 76
Table 4.28 Predictive abilities of models created using only subsets of variables. 77
Table 4.29 Predictive models produced to model the occurrence of *Phalaris arundinacea* and *Petasites hybridus*. 80
Table 5.1 Other work on the occurrence/spread of the study species in habitats. 89
Table 5.2 Species characteristic of the upper-riparian bank habitat. 90
Table 5.3 Vegetation and substrate categories recorded during the fine scale studies. 91
Table 5.4 DAFOR abundance recording methodology. 91
Table 5.5 Density recording categories. 91
Table 5.6 Correlations between proportional species occurrence along bank zones of the Wear. 102
Table 5.7 Autocorrelation of species' occurrences with distance in different bank height zones. 103
Table 5.8 Correlations of proportional species occurrence with RCS variables. 104
Table 5.9 Predictions of the proportional occurrence of *Impatiens* in the lower bank zone of the River Wear 107
Table 5.10 Autocorrelation of species with distance along the river (using 50m data). 109
Table 5.11 Regression models using species data at 50m resolution with the 50m habitat data and selected RCS variables. 110
Table 5.12 Variables in the logistic regression models summarised in Table 5.11. 110
Table 6.1 Density dependent mortality of species from amalgamated monthly data. ..131
Table 6.2 Density dependent mortality of species using individual monthly data. 132

Table 6.3 Seed estimates for Heracleum plants from the literature 141
Table 7.1 Variation in altitude, climate etc. of the altitudinal study sites. 147
Table 7.2 Seed-head characteristics of the flowering *Heracleum* plant at Wolsingham.171
Table 7.3 Windspeed and precipitation data for 1997 and 1998.....180

Chapter One

Aims and Background

1.1 Invasions and the study of non-indigenous species

1.1.1 What is a non-indigenous species?

Non-indigenous species, hereafter referred to as aliens for brevity, can be defined as any species occurring in a defined geographical area in which they are not native, as a result of passive or active introduction by humans (Cowie & Werner 1993). The scale of such an area is somewhat arbitrary but is often set at the level of a country (as in many biological floras). Invasive aliens are defined as those aliens capable of reproducing by any means and establishing new individuals without deliberate human assistance and past the first generation outside of cultivation (Cowie and Werner 1993). This definition includes colonists - species that are unintentionally introduced, and casuals - non-native species that only persist for a short time, as aliens. Alien species can be further divided into neophytes, i.e. newly introduced species (often taken as species introduced after about 1550 AD, e.g. Webb 1985), or archaeophytes, which hail from more ancient introductions.

1.1.2 An introduction to the problems associated with alien invasions.

Non-indigenous species, whether plant or animal, are now a significant component of ecosystems the world over (Heywood 1989), though some habitats remain comparatively free from aliens e.g. tundra and boreal forest. Such species, whether introduced deliberately or accidentally, have led to various problems in different localities (Macdonald *et al.* 1989). Introduced fish species in the African great-lakes have devastated native fish populations (Payne 1986); rabbits have led to extensive over-grazing in areas of Australasia (Myers 1986); rats and goats introduced to islands have led to the loss of endemic species (Atkinson 1985; Coblentz 1978); large areas of natural vegetation of South Africa have been replaced by alien shrubs (Heywood 1989). The potential threat of new "genetic aliens" arising either from hybridisation of species with genetically modified organisms (GMOs) or from the naturalisation of GMOs themselves may similarly prove problematic for native species the world over in the future (Butler and Reichhardt 1999; ACRE 1999).

One major problem with studying invasive species is that almost invariably such species are still in a process of further spread. This fact hinders attempts to predict species

spread into new areas and characterisation of species' habitat traits and demographic limits (Beerling 1993; Pysek *et al.* 1998; Sindel & Michael 1992). Ideally experimental introductions into unoccupied areas would assist predictions but such methods are unethical (and often illegal). A partial solution is to study the species in controlled environments (Beerling 1993) or in its native environment and interpret results accordingly but there is still a large element of uncertainty in predicting the probability of invasion into new areas.

Invasive alien species are widely reported as causing problems in the regions into which they are introduced, largely through the disruption of original ecosystem functioning (Macdonald *et al.* 1989; Vitousek 1992; Broembsen 1989). Problems often associated with alien invasions include a reduction in diversity of native species (Rubec & Lee 1997; Jensen 1976; Loope 1992; Smith 1985) and economic problems arising from the disruption of natural resources or colonisation of agricultural lands (e.g. Payne 1986; DeBach 1974; Lattin & Oman 1983).

1.1.3 Background to the study of alien plant species

There is a mass of literature available describing plant invasions, attempts at managing such species and means of predicting which species will become invasive. To review this literature or even solely that of riparian alien invasions would prove too great a task to undertake here. It is appropriate however to provide a brief resume of research on aliens to give the reader a feel for the extent and type of work that has been undertaken elsewhere.

The study of alien species is by no means a new field, having been a subject of interest to biologists for many years (e.g. Pliney AD77; Ray 1686; Darwin 1872; Marsh 1865 and more recently Elton 1958 and Salisbury 1961). Alien invasions probably started in earnest in the period of recorded history (di Castri 1989; Kornas 1983). More recently however this field of biology has blossomed, largely as a result of the increased rate of introductions between the old and the new world which began in the Victorian era of long distance trading of goods (di Castri 1989).

Reviews of alien invasions have been published for many areas of the world including: Australia (Groves & Burdon 1986), North America and Hawaii (Mooney and Drake 1986; Stone *et al.* 1992, Devine 1998), South Africa (Macdonald *et al.* 1986), Mediterranean ecosystems (Groves and di Castri 1991), the British Isles and Europe (Clements and Foster 1994, di Castri *et al.* 1990), regions of the tropics (Ramakrishnan 1991) and also the marine environment (Jansson 1994, Boudouresque *et al.* 1994). The

main region for which scant information is available, at least in English, is that of central and east Asia. In addition to texts dealing with invasions in biogeographic regions there are also reviews that deal with invasions as a global phenomenon, e.g. Drake *et al.* 1989, Lever 1994. These reviews highlight the fact that invasions by alien plant species are a global phenomenon and that such invasions often have severe ecological and economical impacts.

1.1.4 Approaches to the study of alien plant invasions

There have been many attempts to predict which traits will lead to species becoming invasive (Simberloff 1989; Brown 1989, Thompson, Hodgson & Rich 1995) or to describe the ideal invader (Baker 1965; King 1966; Holm *et al.* 1977; Elmore and Paul 1983). In a similar vein people have categorised invasive species to produce lists of the most prevalent species (Holm *et al.* 1977, Crawley 1987), or alternatively used the approach of quantifying the invasibility of different habitat types (Pysek 1994; Pimm 1989; Rejmanek 1989; Brown 1989; Crawley 1987; Newsome and Noble 1986; Baker 1986).

Predictions of the distribution of both native and alien species have either used static models based on climate/habitat factors (e.g. Pysek *et al.* 1998; Narumalani *et al* 1997; Beerling *et al.* 1995; Huntley *et al.*1995, 1989; Beerling 1993; Prentice *et al.* 1992), or dynamic models incorporating temporal interactions (e.g. Collingham *et al.* 1996; Colasanti & Grime 1993; Allen *et al.* 1991; Hanson *et al.* 1990; van den Bosch *et al.* 1990, Reeves & Usher 1989). Little attempt to combine the two has occurred until recently (Hengeveld 1994, though see Collingham *et al.* in press). Few models have yet proved capable of explicitly and mechanistically modelling the rare long-distance dispersal events that are such an important component of alien invasions (Green & Johnson 1995) and indeed all large-scale species range changes.

There are only a few integrated studies that have incorporated species-habitat interactions in the modelling of alien invasions (e.g. Pysek 1994) and few, if any, have utilised demographic variation between habitats. The work presented here, examining the effects of climate and habitat on species demography, with the aim of both substantiating predictive models and potentially parameterising dynamic models, developed by others in the research group, is therefore largely unique.

1.2 Aliens in the British flora

It has been estimated that over 40% of the British flora is now made up of alien plant species (Clement & Foster 1994). Of these several species have spread rapidly to

become problematic in various artificial and natural habitats: Crawley (1987) provides a list of the twenty most invasive British aliens. It has also been found that particular habitats in Britain, such as woodlands, rivers and wasteland are more prone to invasion than others (Crawley 1987). With such factors in mind the work of this thesis studied the occurrence of three of the more invasive alien plant species along watercourses, thereby incorporating the most invasive species and also the most invasible habitats.

1.2.1 Why are river systems so invasible?

There are several reasons why riverine systems may be prone to alien invasions:

- mesic habitats have been shown to carry larger proportions of alien species than other areas (Rejmanek 1989)
- much of their vulnerability can be ascribed to the availability of "safe sites" that are free of environmental hazards (Harper 1977) or which have no environmental resistance (Johnstone 1986). Simberloff (1986) suggests a similar mechanism with his biotic resistance hypothesis.
- Other hypotheses that have been put forward (Ashton and Mitchell 1989) as making systems susceptible include:

i) species-poor communities (Elton 1958),
ii) poorly adapted native species (Sculthorpe 1967),
iii) absence of predators (Mitchell 1974; Harper 1977),
iv) gaps generated by disturbance (Mitchell 1974, Sousa 1984),
v) competitive superiority due to tolerance of lower resource levels (Mitchell 1974, Connell and Slatyer 1977; Noble and Slatyer 1980) and
vi) empty niches (Elton 1958; Harper 1977; Cook 1985; Johnstone 1986).

All of these hypotheses can be applied to the riparian environments, which are prone to winter flooding and scouring, leaving large open areas. Similarly the nutrient rich substrates that often occur along river margins provide an ideal medium for opportunist alien species (Hobbs 1989). Additionally the concentration of urban areas along waterways provide not only a source of continual system disturbance but also of alien propagules, hence increasing the number and timing of introductions; factors that Crawley (1989) emphasises as important in the invasion process. Finally the river itself can provide a rapid means by which plant propagules can spread large distances over short time periods (Wadsworth *et al.* in press; Nilsson *et al.* 1991).

Many riparian environments have been shown to have a high proportion of alien plant colonists: Decamps and Tabacchi (1994) recorded areas in France where over 25% of

all riparian species were aliens; Dixon and Johnson (1999) reported 33% of all plant species of an Idaho river as exotics; Ellenberg (1988) describes alien species as becoming an increasing component of central European riparian habitats. As rivers reach the coast this situation is further exacerbated by colonists from ports and industrial hinterlands (Crawley 1989a); for example sixty-six species of alien were recorded from Irish coastal ports (Reynolds 1996). The establishment of species as aliens as a result of ships emptying ballast at dock is an important source of riparian aliens (Baker 1986, Graham 1988, National Research Council 1996).

1.3 The study of riparian invaders – The study species

The three alien species chosen as the study species for this research are *Impatiens glandulifera*† (Royle), *Heracleum mantegazzianum* (Sommier & Levier) and *Fallopia japonica* (Houtt.). These species names are simplified to *Impatiens, Heracleum* and *Fallopia* respectively through the remainder of the thesis, except when mentioned in conjunction with species of the same genus at which point full Latin names are used. The three species make for a good comparative study as they all have quite different life strategies, which are detailed in the following sections.

1.3.1 Brief history & life history of the three species

All three study-species were originally introduced into the British Isles as garden plants during the last century (details below). From these original introductions they subsequently escaped and naturalised. The species are still consolidating their ranges, though they all now occur through most counties of the British Isles. The three species are also noted as widespread alien species through much of Europe, *Heracleum* and *Fallopia* also having naturalised in North America and Canada (Beerling & Perrins 1993; Beerling *et al.* 1994; Tiley *et al.* 1996).

1.3.2 *Impatiens*

Impatiens originates from the Himalayas, occurring from Kashmir to Garhwal between 200-2500m and is frequent along roadside ditches and around field borders (Gupta 1989). The first authenticated record for *Impatiens* in the British Isles was in 1839 (Coombe 1956). The subsequent spread of *Impatiens* has been well documented (Trewick & Wade 1986; Beerling & Perrins 1993, Perrins *et al.* 1993), the current distribution (Figure 1.1) showing the species to be widespread over most of England and occurring less commonly throughout Scotland and Ireland.

† Vascular plant nomenclature follows Stace (1991).

Impatiens is an annual species, the tallest in the current British flora (Grime *et al.* 1988), which is expanding predominantly in riparian habitats (Pysek and Prach 1994, Trewick and Wade 1986). Germination of its entire viable seed population occurs in the spring (Grime *et al.* 1988), supposedly with no dormant seed bank though this requires clarification (Beerling & Perrins (1993) give an example). Flowering occurs from late June until the first autumn frosts, producing pods from which seeds are dispersed by a ballistic strategy to over-winter, either *in situ* or after further dispersal by winter floods. Pollination of flowers is quite specialised, predominantly utilising widespread *Bombus* spp. (Dunn 1977).

1.3.3 *Heracleum*

Heracleum is native to the Caucasus, south-west Asia (Mandenova 1950) where it occurs in forest edges and glades, often at stream sides in montane areas with annual rainfall between 1000 and 2000mm (Pysek 1991; Ochsmann 1992). The first record of *Heracleum* in the British Isles was in 1828 (Perring *et al.* 1964), its subsequent spread being documented by Tiley *et al.* 1996. The current species' distribution (Figure 1.1) shows a widespread but rather sparse occurrence over the British Isles, with the exception of the south-east and some areas of Scotland.

Heracleum is the tallest monocarpic perennial in the current British flora, growing up to 5m in height. Flowering does not occur in the initial growing season following germination but most often in the third or fourth season (Tiley *et al.* 1996). Flowers are non-specialised and can be pollinated by many insect species (Grace and Nelson 1981).

The species is most frequently encountered as a colonist of open ruderal communities, disturbed habitats or bare ground, especially in riparian habitats (e.g. Pysek 1994). However secondary spread can occur onto several other habitat types (Tiley *et al.* 1996). Mature plants produce vast amounts of seed (sometimes over 100,000, (Tiley *et al.* 1996)), which are initially wind dispersed but which may subsequently be further spread along waterways. The frost-hardy seeds usually germinate in January-March, plants growing rapidly through the growing season before shoots senesce in the autumn; a similar pattern occurring in subsequent seasons until flowering occurs. Knowledge of the species seed bank is by no means complete though seeds are reported to remain viable for between 7 years (Vogt Anderson & Calov 1996) and 15 years (Lundstrom 1989).

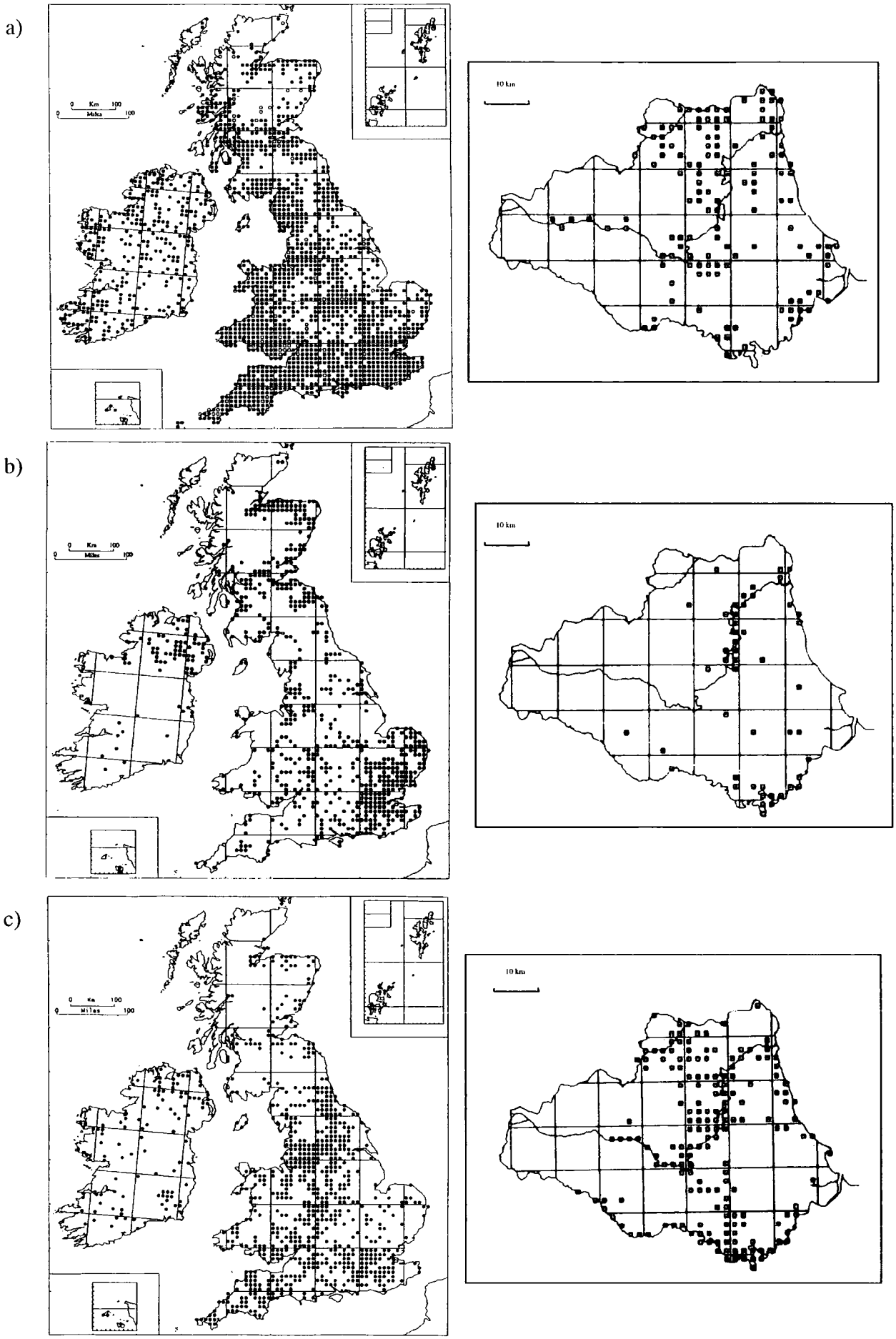

Figure 1.1 Distribution of a) *Fallopia* (1995), b) *Heracleum* (1995) and c) *Impatiens* (1992) in the British Isles (left) and Durham (all 1988) (right). British maps from the Biological Records Centre and Durham maps from Graham (1988).

1.3.4 *Fallopia*

Fallopia is native to China, Japan, parts of Korea and Taiwan (see Beerling *et al.* 1995 for its native distribution) but was introduced into the British Isles around 1825 (Conolly 1977). The species current British distribution (Figure 1.1) depicts a widespread but patchy distribution with almost constant occurrence (at a hectad scale) over much of the south-east, south-west, Wales, Cheshire, Cumbria, Durham and the Scottish Borders, with more scattered localities elsewhere. Despite its reliance on vegetative dispersal *Fallopia* is more widely distributed than the other two species.

Fallopia japonica is the tallest polycarpic perennial herb growing wild in the British Isles, although applying the term polycarpic is something of a misnomer as the species never ripens seeds in Britain, although seeds resulting from hybridisation with *Fallopia baldschuanica* have occurred (Bailey 1988; Bailey & Conolly 1991). Instead it depends entirely on vegetative expansion and fragmentation for its spread. In this respect it is probably more dependent than the other two species on anthropogenic long-distance dispersal. Plants occur mainly in habitats strongly influenced by humans, often relying on transport in soil or as garden waste to establish in new areas. In the riparian environment the species is capable of spreading more naturally due to fragmentation and wash-out of plant parts during flood events. Shoots grow from rhizome fragments, which can be as small as 8g (Beerling 1990a), and during the growing season some resources are translocated to underground rhizomes (Beerling *et al.* 1994), the above ground shoots senescing before winter.

1.3.5. The invasive problems of the study species with particular reference to their spread in the British Isles.

There are several problems associated with the spread of the study species in the British Isles. One of the most often cited threats from alien species, a loss of native species following colonisation, has been linked with these species (Palmer 1994; Tiley & Philp 1994; Vogt Andersen 1994; Williamson & Forbes 1992; Jensen 1976). Another problem frequently associated with all three species is a loss of ground flora beneath stands, which leaves riverbanks more susceptible to flood erosion in the winter (Beerling 1990b; Tiley & Philp 1994; Roblin 1994). *Heracleum* poses a threat to public health due to its production of furanocoumarins, which can lead to photodermatitis after contact with bare skin, often leading to serious injuries (Tiley & Philp 1994; Camm *et al* 1976; Morton 1978; Rubow 1979; Lundstrom 1984; Wyse-Jackson 1989; Drever & Hunter 1970). There is the possibility of a threat to livestock from *Heracleum* (Andrews

et al. 1985; Harwood 1985) but this seems a minor problem as many species can happily graze the plant with no detrimental effect (Tiley *et al.* 1996; Vogt Andersen 1994). Other frequently cited problems arising from these species include restricting access to amenity areas (Tiley & Philp 1994; Vogt Andersen 1994; Beerling & Palmer 1994; Palmer 1994), roadside obstructions (Tiley & Philp 1994; Beerling & Palmer 1994; Palmer 1994), crop disease hosts (e.g. *Sclerotinia* sp. on *Heracleum* (Tiley & Philp 1994)) and structural damage to riverbanks, flood defences, pavements and roads (Beerling 1991b; Palmer 1994). Such problems have led to control programmes being adopted in many areas, most of which have achieved only limited success. The work of this thesis, in combination with work of others in the research group, will aid in future planning of control strategies.

1.4 Scope and aims of the thesis

The main aim of the thesis is to initially determine how invasive alien species characteristic of riparian environments are dependent upon different habitats and consequently if species distributions can be modelled using such attributes. The importance of both habitat and climate as limiting factors, and their relative importance in alien invasions compared to other potentially important factors such as dispersal and pollinator limitations was of interest. By selecting three several alien species, with different life-forms and life-history characteristics, it would be possible to determine which features of riparian environments allow such species to invade and whether these species select similar riparian habitat types and are restricted in their occurrence by the same parameters. Fine-scale distributions of the species are also examined to determine patterning in the species abundance distributions throughout their local ranges.

A further component of the thesis examines the effects of climate, as represented by elevation, on the distribution of the three species, to elucidate whether the species currently fill all zones of suitable climate or whether dispersal or other factors are limiting further spread. This climatic work also utilises demographic traits to investigate if particular stages of growth are responsible for limiting further spread. Work such as this, elucidating mechanisms operating to limit plant species has been emphasised as necessary in order to place confidence in bioclimatic predictions (Panetta & Mitchell 1991; Pysek *et al.* 1998, Beerling 1993). By answering these questions and discussing the results in light of other work, this thesis highlights factors limiting species spread at local and catchment scales and helps answer the question of the importance of scale in species modelling.

This work forms one component of a larger project, examining spatial and temporal scale dependencies in the invasion of riparian habitats by alien weeds. The aim of the project was to model the distribution of several alien species at various scales, taking account of dispersal attributes, climatic variables and land-use characteristics. The work presented in this thesis is primarily working at smaller temporal and spatial scales. Nevertheless the thesis will be presented in relation to the modelling of alien species at different scales and will highlight contributions to general ecological theory and conservation practice.

1.5 Plan of the Thesis

Chapter One is split into two sections. The first details the background to previous work undertaken on invasive alien species and the second (above) describes the scope and aims of the thesis. The following chapter then goes on to describe the study area in terms of climate, geology and habitats.

Chapter Three initially describes the River Corridor Survey (RCS), which is used as a tool to derive study species distributions and from which habitat and land-use data were derived. Following this the methodology of data extraction are described and, following some simplification and amalgamation of data, initial exploratory statistical analyses of the data in relation to the study species are undertaken. Characterisation of the catchments from the RCS data and study species distributions are presented, along with a discussion describing the use of the data for subsequent modelling and limitations of the data.

Chapter Four provides a brief overview of previous modelling of riparian species distributions. Techniques of modifying the data so as to limit effects of autocorrelation are implemented before introducing logistic regression modelling as a means of predicting species distributions using RCS data. Resultant predictive models are tested both within and between catchments and the results discussed in the light of known species ecology.

In Chapter Five refined scales of both species and habitat data are introduced, collecting data not only at a smaller spatial scale than the RCS but also collecting data from different height zones of the riverbanks. The distributions of the alien species at these fine scales are examined to detect similarities and differences in distribution and abundance patterns. Finally the chapter investigates whether such fine scale data result

in better predictive modelling capabilities and introduces modelling using quantitative data. The results are compared to those produced in the previous chapter.

The sixth chapter builds on the findings of the previous two chapters by examining the effects of habitat type on plant demography, with the aim of explaining why species habitat selection occurs and factors likely to limit or encourage further spread. The germination, growing season and over-winter mortality, and productivity of the species are examined, where possible, in different riparian habitat types. The results are discussed in relation to published performance information for the three species and in light of modelled habitat selection.

Chapter Seven introduces climate as a variable in the prediction of alien species occurrence and spread. An experimental transect along an altitudinal gradient is described from which comparative measures of species growth, mortality and performance were collected and analysed. These values were then related to climate using interpolated climatic values for each site. By combining the data from the sites over the two years of study it was possible to look for trends in species performance in relation to climatic variables. The findings are compared to published literature and their potential implications in terms of future species spread are discussed.

The final chapter provides a brief overview of the findings of the thesis and discusses them, both in relation to the published literature and also with regards to future spread of the species and control measures. The results are discussed in relation to a more general overview of species modelling and alien species, and areas of further work are highlighted.

Chapter Two

Introduction to the Study Area

2.1 Climate, geology and habitats

The county of Durham referred to herein includes the districts of Tyne and Wear and Cleveland, and comprises an area approximately 70km East-West and 50km North-South at its greatest extent. It is situated in the narrow neck of northern England with the Pennine fells forming its western boundary and the North Sea its eastern (see Figure 2.1). The boundaries of the county proper are delimited to the north by the rivers of the Tyne and Derwent, and to the south by the Tees watershed and the Tees itself. These river systems along with the river Wear, which meanders through the centre of the county, drain the entire county. To the north it is bordered by Northumberland, to the south by North Yorkshire, and to the west by Cumbria. The highest point in the county, Burnhope Seat, lies in the Pennine uplands at 746m, close to the western border. The ground rapidly falls away to the lowlands in the east, with only one or two plateaux of higher ground away from the Pennines.

The lowlands of the county are quite urbanised and densely populated, especially along the courses of the major rivers and contrast starkly to the relatively unpopulated uplands whence the rivers originate.

The geology of Durham has been well described (Dunham 1948; Smith & Francis 1967; Johnston *et al.* 1970) so only the briefest of mentions is afforded here (adapted entirely from Graham 1988).

Much of the county is covered with peat or glacial drift but the gentle eastward dip of the underlying rocks causes some outcropping in the west. The western fells are comprised almost entirely of Carboniferous rocks, with a limestone-sandstone-grit layering below the Millstone Grit outcropping almost continuously in Teesdale and Weardale (the two major upper valleys). These rocks are almost entirely covered by glacial clay and peat and the limestone only occasionally breaks through to form extensive scars, on which calcicolous species can thrive, much of remaining upland being extensive heather moor. The intrusion of the Whin Sill into an area of limestone baked it to a saccharoidal marble, which has weathered in small areas of the uplands to produce conditions that support unique assemblages of plant species (Ratcliffe 1977; Pigott 1956; Bradshaw 1970; Willis 1995).

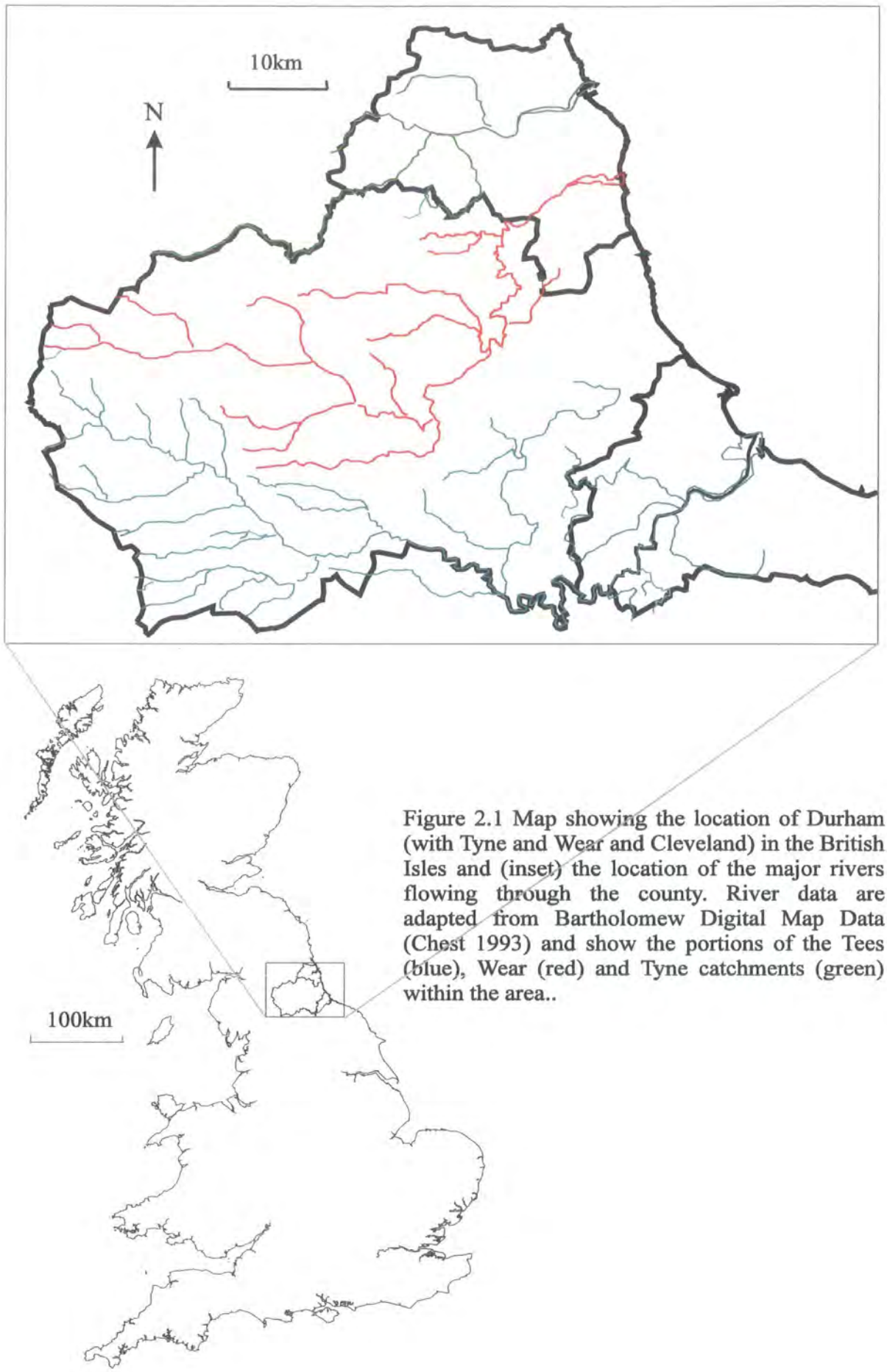

Figure 2.1 Map showing the location of Durham (with Tyne and Wear and Cleveland) in the British Isles and (inset) the location of the major rivers flowing through the county. River data are adapted from Bartholomew Digital Map Data (Chest 1993) and show the portions of the Tees (blue), Wear (red) and Tyne catchments (green) within the area..

Coal measures occupy much of central Durham, which at one time was an area of much mining activity. In the lowlands red sandstone rock dominates beneath the richer soils and is only displaced over the East Durham Plateau, where Magnesian limestone outcrops dominate the geology.

The county plays host to a wide array of plant communities and many habitat types. Much of county's lowlands have suffered from the effects of industrialisation and intensive agriculture, leading to estuarine and wetland areas being reduced to a fraction of their former extent with an associated loss of specialised species in most lowland habitats. The uplands appear to have suffered slightly less from such effects and areas persist which have escaped much of the ravages of "improvement"; though past management has favoured the predominance of sheep-walks and grouse moors at the expense of natural woody and heathy communities. Drainage, fertilisation and over-grazing are more recent threats to the present upland habitats.

The county is relatively poorly represented by aquatic habitats and has a low proportion of woodland compared to elsewhere in Britain. Conversely it has a wide range of upland habitats and retains some of the best remaining hay-meadows in England. Plant species in the county hail from a wide range of biogeographic regions (*sensu* Ratcliffe 1968; Birks 1973; Jermy and Crabbe 1978); the majority of the species coming from the wide element but with many species of the sub-Atlantic, Oceanic and Continental elements, and even Mediterranean, northern-montane, Arctic-alpine, alpine and endemic elements.

In a global context the climate of the county is equable with Britain as a whole, with an absence of extremes of either heat or cold (Smith 1970). The major climate gradient lies in an east-west direction, with a general inverse relationship between temperature and rainfall as one descends to the coast. The wind across the county blows predominantly from a south-westerly direction, with the high Pennines experiencing mean wind velocities double that of the lowlands (save the coastal regions). In the lowlands, local summer winds are usually associated with the sea-breeze effects, which can penetrate several kilometres inland along the Tees and Wear valleys (Smith 1970).

Relatively low mean annual temperatures result from the cumulative effects of northerly latitude, coolness of the adjacent North Sea and the deterioration of climate with altitude to the west. Manley (1936, 1942, 1943) was able to show that temperature maxima at upland sites were comparable to those at Durham City in the lowlands. Temperature extremes are restricted throughout the county but this effect is most pronounced at the coast. The growing season (using a 6.1^{0}C limit) lasts from early April to mid November

in the lowlands, though at altitudes of 500m this is reduced to 5½ months (Smith 1970). There is some variability about this altitudinal trend however, katabatic drainage from summits can invert temperatures at lower altitudes and frost hollows are well developed in places; one of the best documented frost pockets in Britain occurs in the Wear valley close to Durham City, where temperatures can be more than 8°C lower than at a nearby observatory (Smith 1967, Manley 1941, Catchpole 1963).

Much of Durham benefits from reduced precipitation as a result of the sheltering effects of the Pennine barrier, the lowest part of the region is amongst the driest areas of Britain. Even so high levels of precipitation (up to 2000mm an^{-1}) are estimated for the upper slopes of the county (see Figure 7.1 for interpolated precipitation). The lowest precipitation generally occurs in February and March; highest values occurring usually from August to October. The regular incidence of precipitation generally precludes the possibility of prolonged drought.

The North Sea "haar" (coastal fog), one of the most distinctive features of the Durham climate, acts to reduce the hours of bright sunshine over the county considerably, sunshine hours being consistently lower than areas at the same latitude on the west coast of the country (Table 2.1).

Table 2.1 Sunshine hours during haar conditions on 15th June 1967 and monthly means sunshine hours at east and west coast sites (from Smith 1970).

Station	Sunshine Duration (hours)	
	15th June	Monthly Mean
Hartlepool	2.6	5.9
Durham Observatory	6.1	6.0
Dumfries	11.9	6.1
Morecambe	12.4	6.9

Values are shown for Hartlepool at the mouth of the Tees, the Durham Observatory 18km inland on the Wear, Dumfries on the Scottish south-west coast and Morecambe on the west coast of Lancashire (at a similar latitude to the County Durham sites).

Annual sunshine hours show a slight increase inland from the coast and then a decline westwards, with an average of 3.6 hours per day at Durham dropping to less than 2.5 hours on Dun Fell (Smith 1970).

2.2 Durham's alien flora

Within County Durham there are many alien plant colonists; Graham (1988) lists about 38% of all the recorded species (excluding microspecies) in the county as aliens and a further 26% of the remaining species as casual, i.e. non-native but never establishing for more than a short period. Aside from the study species other aliens associated primarily

with the rivers and riparian areas of Durham include *Mimulus guttatus, Saponaria officinalis, Hesperis matronalis, Fallopia sachalinensis, Symphytum* x *uplandicum, Epilobium brunnescens* and *Elodea canadensis*. Non-riparian alien species that are widespread in the county in regions other than along waterways are usually associated with waste areas, railway lines etc. and include notable species such as *Buddleja davidii, Sambucus canadensis, Chamomilla suaveolens* and *Senecio squalidus*. Other aliens have prospered in a variety of areas influenced by human activity e.g. species such as *Trifolium hybridum* and *Spergula arvensis* in farm crops and *Veronica filiformis* in lawns. The woodlands of the county support many alien species including species such as *Symphoricarpos albus* and *Doronicum pardalianches* as well as numerous introduced tree species e.g. *Castanea sativa* and *Sorbus aria*. The upland areas are more impervious to aliens, the moors supporting few such species (*Campylopus introflexus* is an exception), although areas of extensive river gravels can produce some of the richest sites for aliens in the county.

2.3 The County Durham rivers

The lowland river sections in which all three species thrive can be categorised, using the classification of Holmes *et al.* (1999), as type A-IV impoverished lowland rivers. The rivers can probably be further grouped into sub-type AIVc, which are predominantly situated in northern England and are recorded in Northumbria and Cumbria as well as into the Scottish Borders. Such rivers are described as upland-origin rivers with impoverished floras, with high altitude sources and increased occurrence of riffle habitats. These areas typically have a high proportion of silty substrates but also retain cobble, boulder and bedrock areas. However this sub-type, unlike a similar type (AIVb) does not have *Impatiens* recorded as a common species, so it may be that the Durham rivers fall somewhere between these two sub-types. Homles *et al.* (1999) describe such AIV-rivers as occurring over areas of soft geology, being predominantly narrow in width and characterised by degradation of the physical environment through land-drainage and flood defence activity as well as sometimes suffering from low-flows and pollution incidents. Mean numbers of species in and beside such rivers are low compared to other lowland rivers, perhaps making them more amenable to alien invasions. Such rivers are widely distributed throughout Britain and typically contain mostly emergent or marginal species with few submerged aquatics. At mid-altitude localities of the Durham rivers, reaches tend to conform to Holmes *et al.* (1999) type B-VI and C-VIII, referring to sandstone, mudstone and hard limestone rivers of Scotland

and Northern England and to oligo-mesotrophic rivers respectively. Such areas of the rivers do not support *Heracleum* but scattered colonies of both *Impatiens* and *Fallopia* can occur in suitable habitat, given the ability to reach such areas.

In the upper areas of Durham the rivers conform to Holmes *et al.* (1999) type X ultra-oligotrophic rivers. Such sites have a great abundance of cobbles, boulders and bedrock, drain base-poor rock and typically flow through areas where blanket bog or acid-heath dominate the catchment upstream. Bryophytes are a major component of the river and bankside flora, whilst the alien study species are absent from these areas. However other aliens such as *Mimulus guttatus* and *Epilobium brunnescens* can survive in areas of these upper rivers.

The macrophytes of the study catchments have been extensively studied and good summaries of vegetation changes over space and time are given in several papers by Whitton and his colleagues (Whitton & Buckmaster 1972; Holmes & Whitton, 1977a, 1977b; Whitton & Dalpra 1968; Whitton *et al.* 1998).

2.4 Vegetation and ecology of the riparian environment

Much work has been undertaken to establish links between the vegetation of the riparian environment and factors such as flood events, drainage regime and catchment geology (Gurnell 1995; Gurnell & Gregory 1987; Dynesius & Nilsson 1994; Whitton *et al.* 1998; Dixon & Johnson 1999; Nilsson & Grelsson 1990). In addition the potential effects of climate change, both short- and long-term may also affect riparian vegetation over time (Beerling 1993, Beerling & Woodward 1994). River characteristics such as frequency and duration of flooding will have important consequences on species persistence in different areas (Nilsson *et al.* 1991; Cavers & Harper 1966). All of these factors in combination with others such as pollution (e.g. Whitton *et al.* 1998), disease (e.g. Gibbs *et al.* 1999) and anthropogenic effects (e.g. Pysek *et al.* 1998) result in a complicated dynamic process operating to structure riverbank vegetation over both space and time. Much previous work on riparian systems however has concentrated solely on changes in the woody species of the bank, ignoring field and ground-layer vegetation (e.g. Harper *et al.* 1997; Toner & Keddy 1997).

2.5 Riverine species & plant communities in County Durham

The rivers of County Durham are varied in nature between their headwaters and their outlets into the North Sea, though the rivers themselves are quite similar in character. The three main rivers of the Tees, the Wear and the Tyne flow through the Pennine

moorland in their upper reaches and are usually coloured from the peat of the upper-catchment. The rivers are generally stony in nature being mostly bordered by cobbles and rocks and containing extensive areas of riffles, with shorter areas of rapids. The vegetation bordering these uplands is variable but often comprises moorland or acid grassland up to the edge of the river with frequent innundation communities of rushes and sedges. Lower down the valleys, as the rivers become wider they are often fringed by cobble banks, giving way to grazed grassland terraces or are fringed by woodland, with species such as *Alnus glutinosa* and *Betula* sp. as well as more widespread broadleaved species such as the alien *Acer pseudoplatanus.* Fringing grasslands are often quite species-rich, with plants such as *Campanula rotundifolia, Centaurea nigra* and *Geranium pratense.*

In the middle and lower reaches of the Durham rivers, the rivers widen, slow and become deeper. The land-use away from areas of habitation is predominantly arable farming with some mixed farmland. There is increased urbanisation of the catchment, which combined with intensive farming practices results in a reduction of water quality. This situation is exacerbated towards the coastal outlets, which are heavily populated and urbanised. The lower reaches of the rivers are frequently sectioned off from surrounding land-use and are therefore left unmanaged. This results in the formation of species-poor tall weed communities.

Diagrammatic representations of the typical hydroseral community of the lowland and upland river margins are shown in Figures 2.2 and 2.3. These highlight the fact that although many of the surrounding vegetation communities are well defined in the NVC classification scheme of Rodwell (1991-1999), those areas of the lower-riparian zone in which the study species prosper fit poorly within this classification. These areas have some parallels with NVC communities, probably most closely matching the MG1 *Arrhenatherum elatius* grassland (*Festuca rubra* sub-community, *Myrrhis odorata* variant) though having an affinity to several of the tall-herb weed communities (OV24, *Urtica dioica-Galium aparine*; OV25, *Urtica dioica-Cirsium arvense*; OV26, *Epilobium hirsutum* communities). It is perhaps as part of this latter grouping that this vegetation belongs; these communities taking advantage of features such as accumulated strand-line debris (Rodwell 1999). They are generally dependent on anthropogenic events including neglect for their development and are especially widespread through the agricultural lowlands (Rodwell 1999), all factors which mirror the conditions beside the region's rivers. Graham provides two suitable noda into which

the *Petasites hybridus* vegetation types fit well, the IN6 flood-plain community and the MG/PH *Petasites hybridus* tall-herb community. *Impatiens* communities are somewhat neglected by Graham, who tends to spread such vegetation between several different groupings depending on other species present.

The dominant *Salix fragilis* fringing woodland of the lowland areas falls best into the NVC W6a community, although the occurrence of large amounts of *Acer pseudoplatanus* and *Fraxinus excelsior* in the canopy along with a ground flora often reminiscent of an MG1 grassland community do not fit well into this classification. This woodland could similarly be described as a variant of a community described by Tansley (1939) as a sessile-oak woodland stream-side and flush community. The vegetation communities of the upland shingle banks also fits poorly into the NVC classification, though the gravel and shingle communities of Graham (1988) exactly match such areas.

The intermediate communities in which the alien species most commonly occur also fall largely outside the scope of other classification systems developed for British river vegetation (e.g. Holmes *et al.* 1999). However, by widening the search for comparable vegetation types it is possible to find analogies for all of the poorly represented habitats at a European level. The *Salix fragilis* woodlands fit well as a variant of with the dominant central-European floodplain white-willow wood (Ellenberg 1988), occurring where the gravels are deficient in lime. Other fringing woodlands fit well into several of Ellenberg's European categories including the Stellario-Alnetum glutinosae and Alno-Fraxinetum stream Ash-Alder woodland. Similarly *Petasites* riverbank communities have been classified by Oberdorfer (1979) and Lohmeyer (1970) as belonging to the alliance of shaded woodland skirts (Aegopodion and Aegopodio-Petasitetum respectively). Ellenberg similarly places these stands in the Aegopodion as Petasites hybridii, classing it as a fringe community requiring nitrate and a high humidity. These classifications make sense given the gradation of *Petasites* into *Aegopodium* along the Durham rivers and help in placing the woodland fringe into a community classification. Moor (1958) even classifies stands of *Impatiens* into an association, the Impatienti-Solidaginetum, which is described as being widespread from Switzerland to Czechoslovakia.

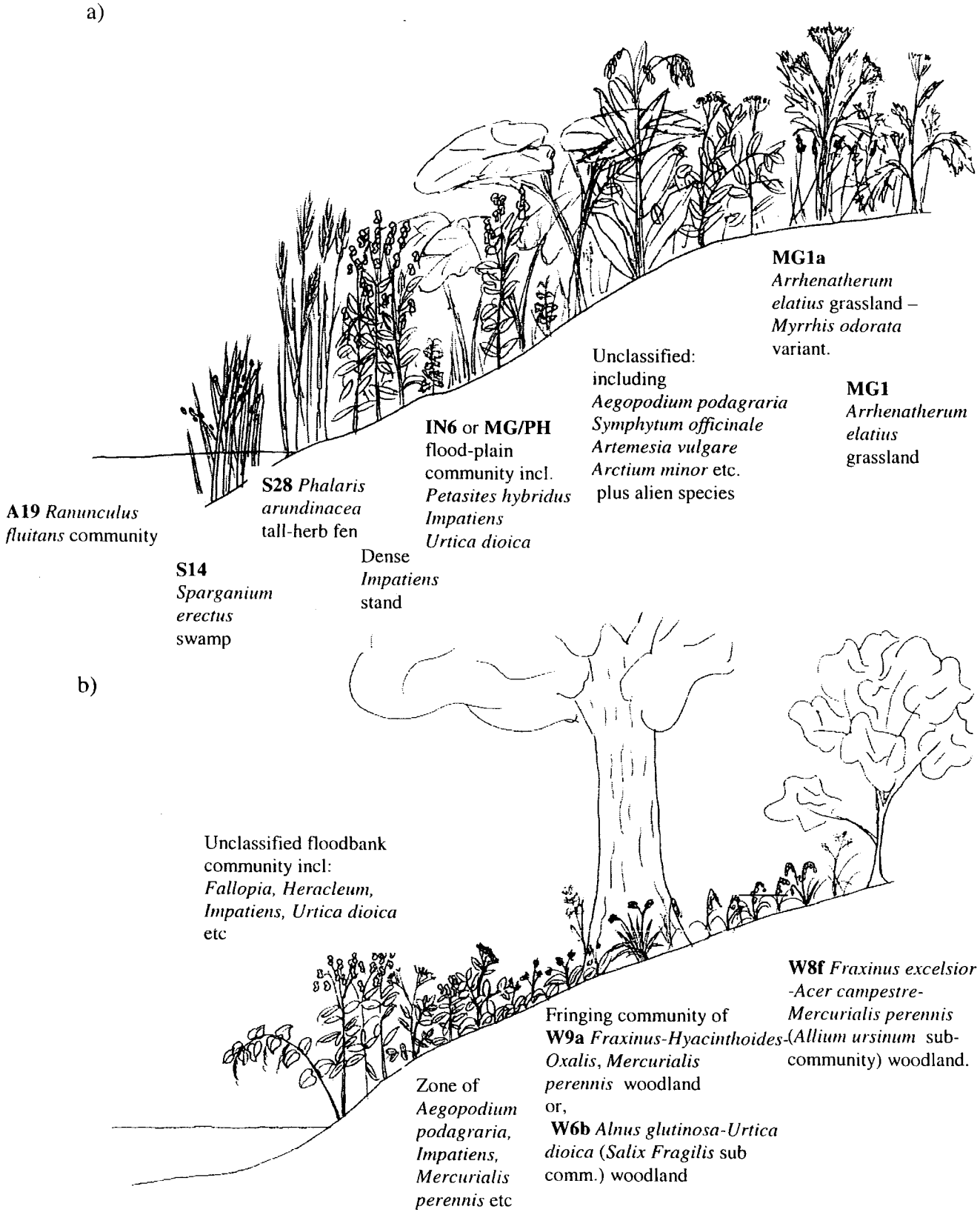

Figure 2.2 Typical cross-sections of the lowland reaches of the larger Durham rivers showing the river abutting a) typical unmanaged bankside and b) woodland. Communities are as described in the NVC classification (Rodwell 1991-1999) or from Graham (1988).

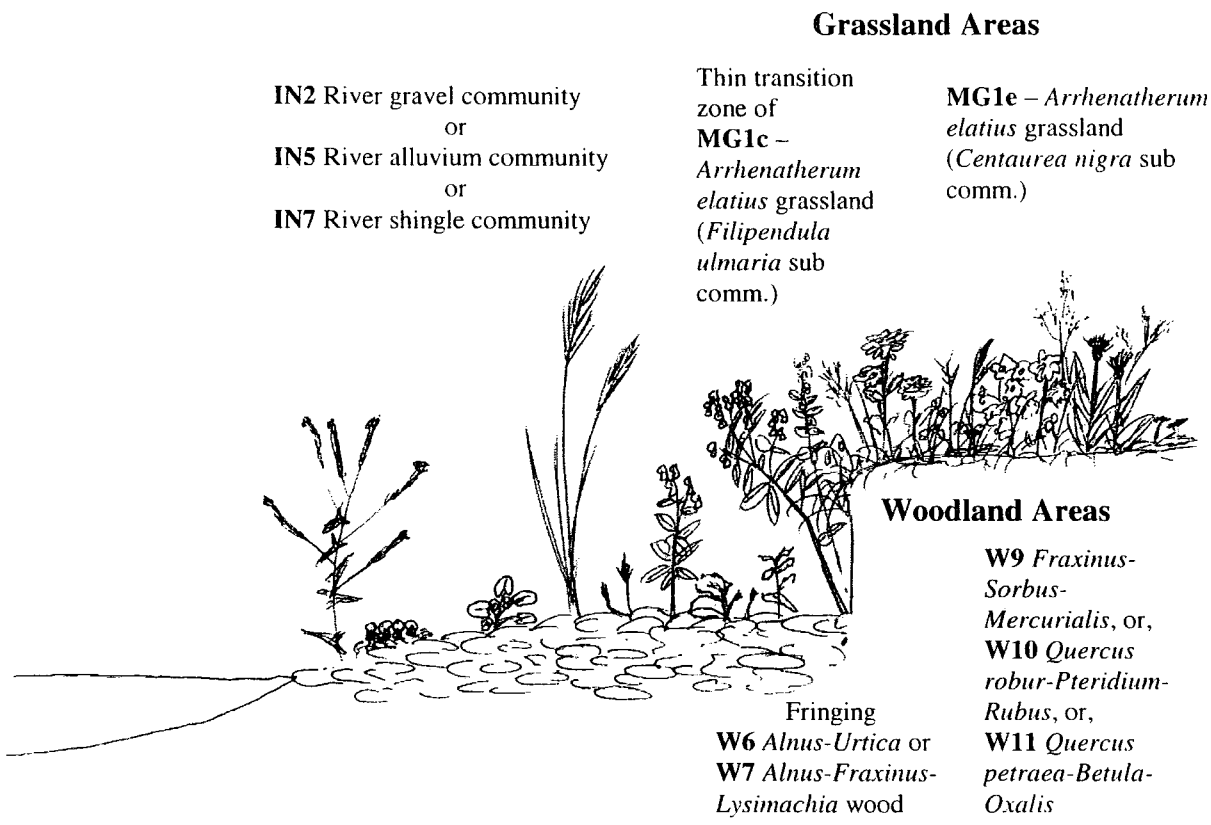

Figure 2.3 Typical cross-sections of the upland reaches of the larger Durham rivers showing the shingle-fringed river abutting typical unmanaged grassland. Also shown are typical adjacent communities for woodland. Communities are as described in the NVC classification (Rodwell 1991-1999) or from Graham (1988).

Chapter Three
The River Corridor Surveys

3.1 Aims and Introduction

The aim of the work presented in this chapter is to examine the distribution of the study species in relation to characteristics of the Wear and Tees river-catchments, by means of analyses of data derived from the Environment Agency's River Corridor Surveys (RCS). This work should determine whether RCS-derived variables relate to the distribution of the study species and whether they can be used to model species occurrence.

Initially the chapter describes the RCS and the methodology adopted to extract meaningful data from it. The resultant data are then processed so as to retain only those variables likely to prove good predictors of species occurrence. Ecologically sensible amalgamated variables are also created to group like-variables that separately may not appear important. Summary and initial exploratory statistics are then undertaken on these variables.

3.1.1 What are the River Corridor Surveys?

The Environment Agency (formerly the National Rivers Authority) is obliged under section 16 of the Water Resources act (1991) to further conservation, to protect sites of conservation interest and to take account of the effects any developmental/management proposals would have. The River Corridor Survey (RCS) is designed to act as a baseline dataset, ensuring special conservation interests of sites associated with the water environment are sustained. The surveys are intended to encompass all of the major rivers and tributaries within a catchment. They are based on recording the major habitats, vegetation and physical features of the river corridor, rather than detailed species or community records. This methodology provides the framework for a consistent national approach for gathering and recording information

3.1.2 RCS data collection methodology

A river corridor is the term used to describe a stretch of river, its banks and the adjacent land (NRA 1992). Usually the river corridor includes land and vegetation within 50m of the river bank and should include four zones: the aquatic zone, the marginal zone, the bank zone and the adjacent land zone.

The methods and procedures of data collection and presentation are set out in detail in the NRA's Conservation Technical Handbook (1) (NRA 1992) so only a brief summary is given here.

The river is divided into contiguous 500m sections, which are surveyed between April and October, recording and mapping the features shown in Table 3.1. Standard recording symbols (described in Appendix I, Figure AI.1) are used on the maps (e.g. Figure 3.2-3.3). As well as a general map of the habitats in a section, the survey reports include river cross-sections, summary descriptions of the general-ecological and morphological character of the corridor and any other points of interest. In the case of the RCS data used here, a record of the occurrence of selected riparian plant species was also part of the final report for each section.

Table 3.1 Features recorded in RCS maps.

Zone	Features
Aquatic Zone	- plant communities
	- flow and current features
	- substrate and physical features
Marginal Zone	- plant communities
	- substrate and physical features
Bank Zone	- tree species
	- other plant communities
	- physical features
Adjacent Land Zone	- habitat types
	- land use

3.1.3 The Tees and Wear catchment RCS

The River Corridor Surveys (RCS) comprise detailed data for each 500m stretch of the main river systems of the Wear and Tees catchments (see Figure 3.1, showing the main rivers of Durham) These data represent biological and physical features (>60 variables); neighbouring land use (25 variables); and the abundance (DAFOR scale: D-dominant, A-abundant, F-frequent, O-occasional, R-rare) of 243 species of macrophyte (Appendix I, Table AI.6). Shortened abbreviations used to represent these variables hereafter are given in Appendix I (Table AI.1). These data comprise 365 and 788 500m stretches for the Wear and Tees catchments respectively. Most of the surveying occurred during 1991-1992 on the Wear catchment (though some areas were surveyed as late as 1994) and 1991-1993 on the Tees catchment (with a few sections surveyed in 1989 and 1990).

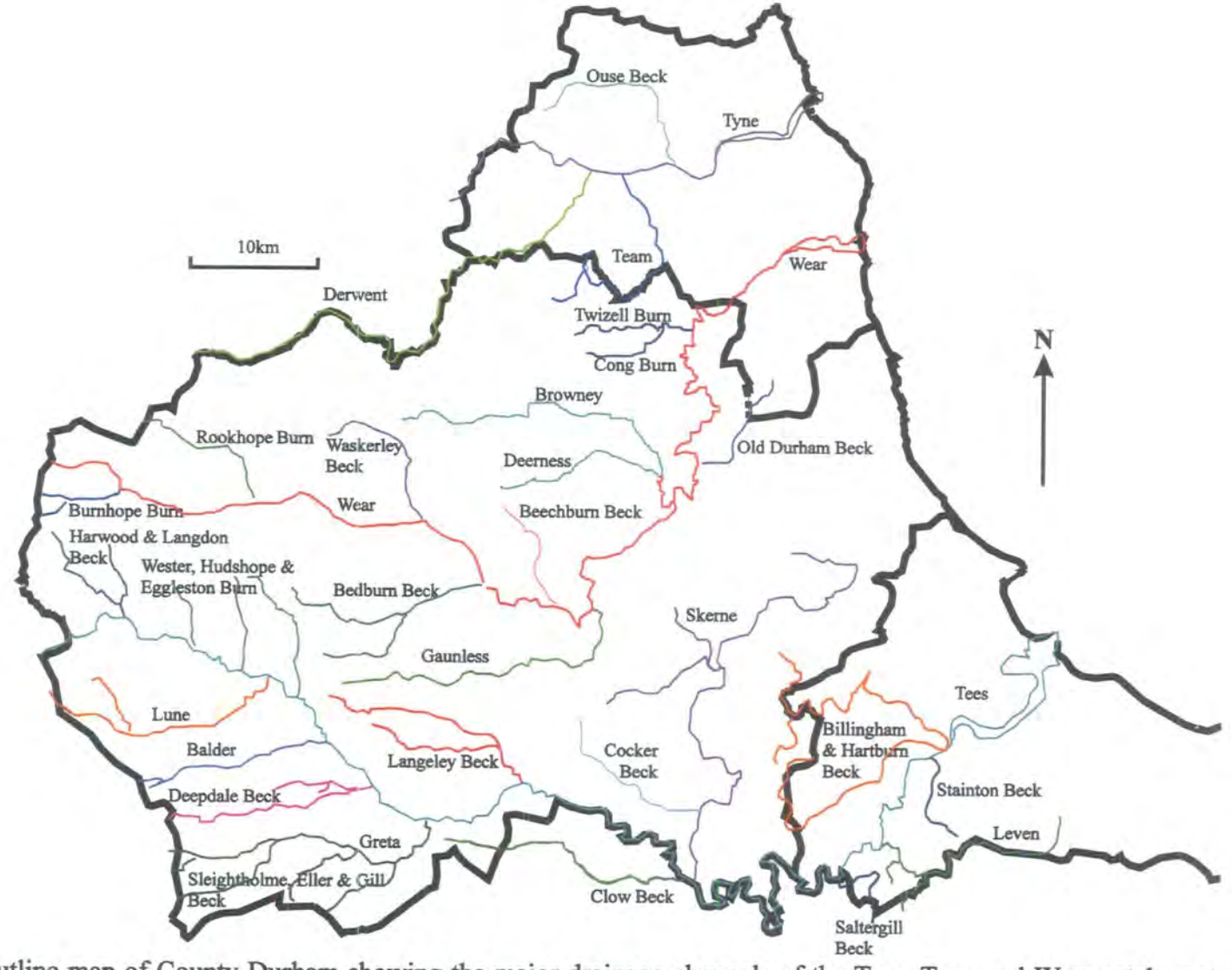

Figure 3.1 Outline map of County Durham showing the major drainage channels of the Tees, Tyne and Wear catchments. Rivers are adapted from the Bartholomew Digital Map Data CHEST (1993).

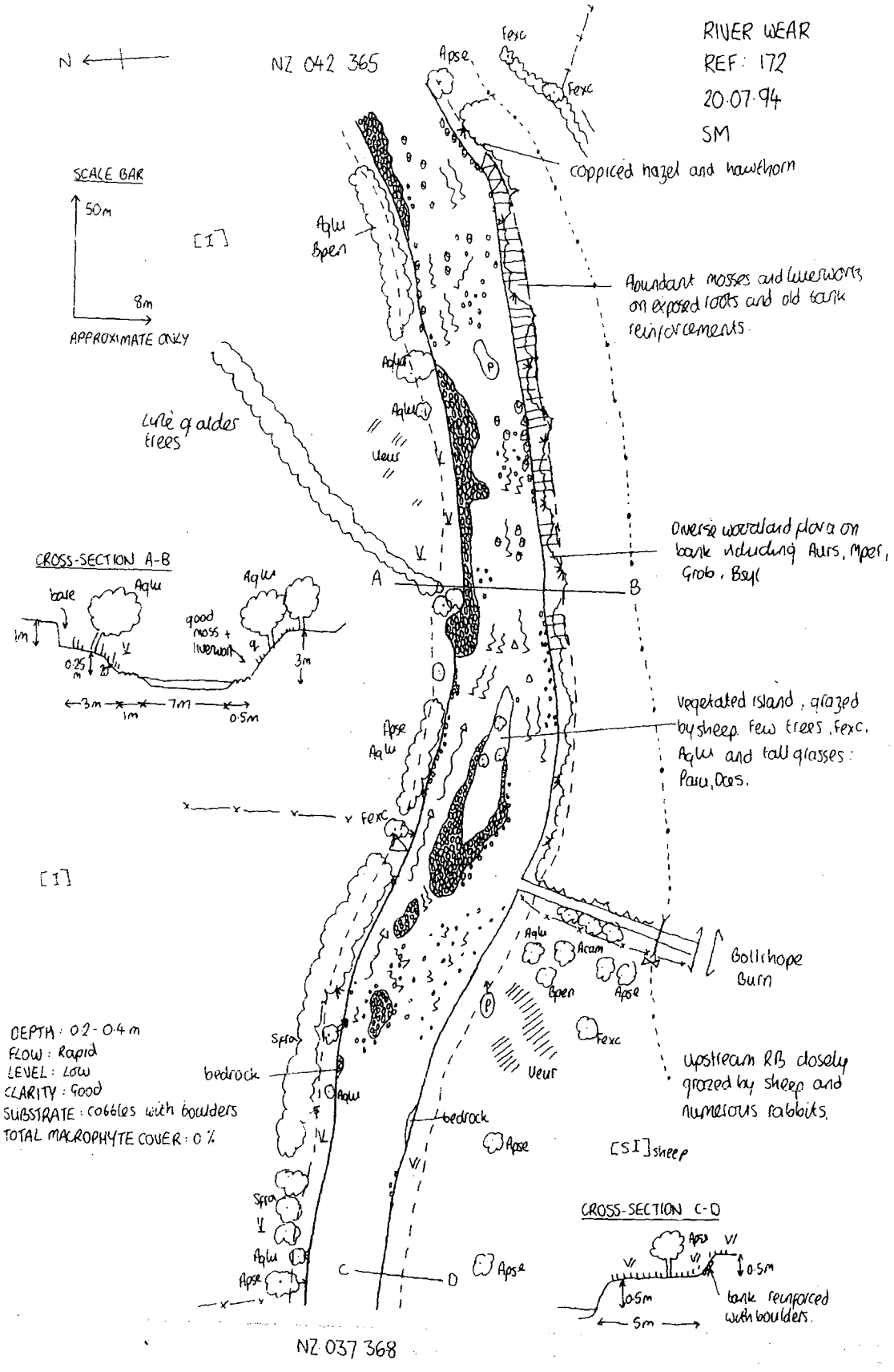

Figure 3.2 Example of a reasonable quality RCS section diagram including much detail on substrate, cross-sections, vegetation types and surrounding land-use. See Appendix I, Table AI1 for symbol explanations.

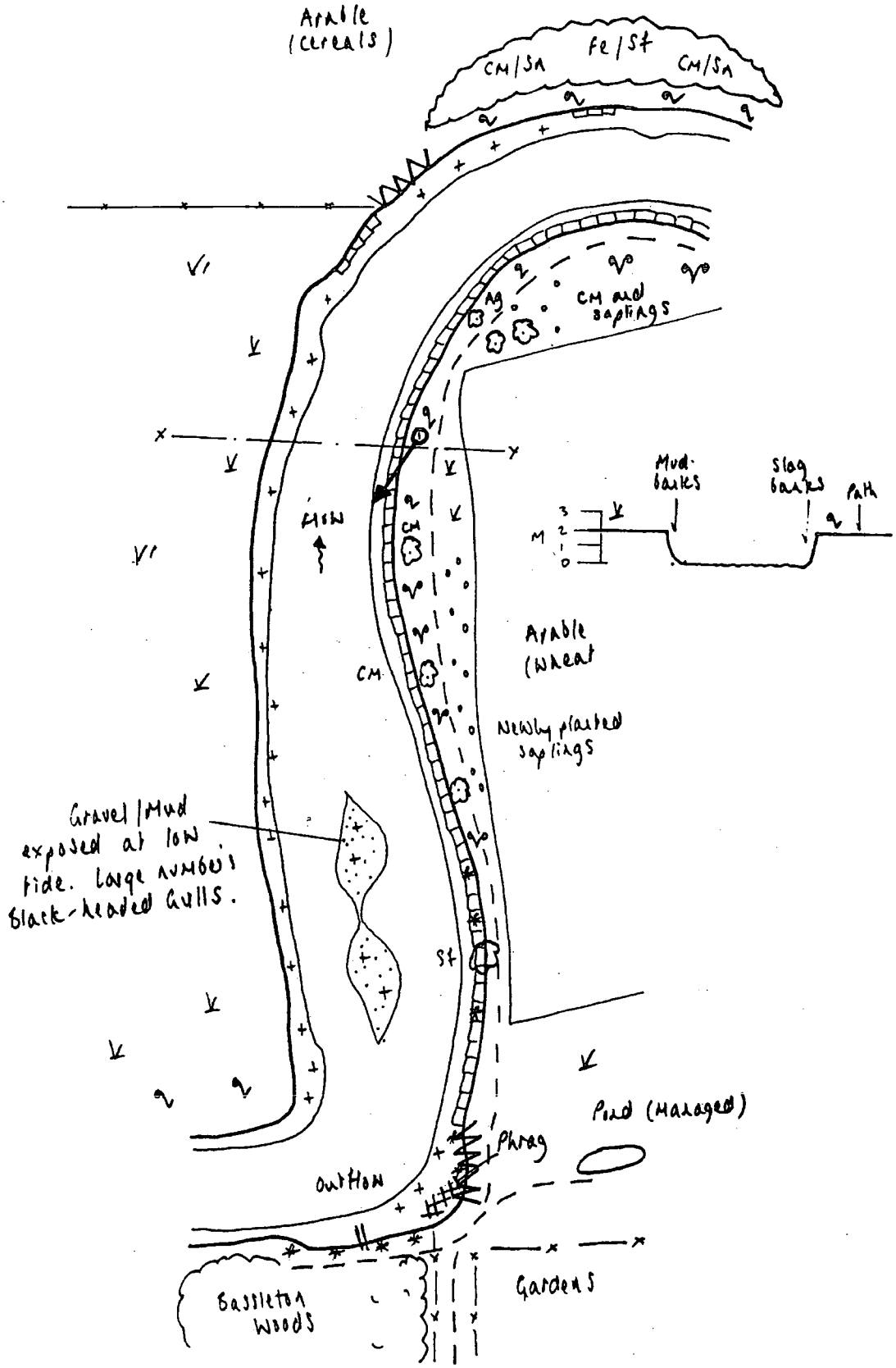

Figure 3.3 Example of a poor quality RCS section diagram. Note the lack of non-arboreal species data. See Appendix I, Table AI1 for symbol explanations.

3.2 Methods

3.2.1 Data extraction methods for the RCS data

In order to examine the distribution and spread of the study species, it was first necessary to input as much currently available data as possible from the RCS. This could then be compared to the occurrences of the alien species also provided in the RCS.

The RCS data were split into left- and right-bank where possible, thereby doubling the number of sections on which analyses could be performed and refining the habitat data with respect to alien occurrence. With such a huge amount of data available to analyse it was necessary to undertake mainly simple descriptive statistics which would aid elucidation of those variables most important in determining distributions.

Details of the database

The information held within the survey sheets is split into sections, all of which were used as information sources. The sections include a summary page, descriptive text on physical/biological properties and channel morphology; a schematic map including information on habitats, substrates and surrounding land-use; and also a macrophyte data-sheet. The macrophyte data are split into bankside and aquatic occurrence and also differentiate between the left- and right-bank.

A list of the summary features that were recorded from the RCS survey sheets is given in Table 3.2. Tables of the other recorded variables are provided in Appendix I, Tables AI.2-AI.5.

Most of the riverine morphology features were recorded simply as presence or absence. The abundance of riffles was entered on a scale from 0 to 5 following a DAFOR scale with 1 representing rare through to 5 for dominant. The abundance of pools within the river was simply the number of pools present in a stretch.

Table 3.2 Basic variables recorded for each section of the River Corridor Survey

Variable	Values
Length Code Number	Bro001, Wea032 etc.
Bank Side	l or r
Downstream Grid Ref.	Six fig. ref.
Upstream Grid Ref.	Six fig. ref.
Height above sea level	Altitude in metres
Date of survey	xx/yy/zz
Minimum Bank Ht (m)	0-30
Maximum Bank Ht (m)	1-40
Minimum Bank Slope	0-90°
Max. Bank Slope	0-90°

Most information derived from summary pages referred to the river in general, not to individual banks, and was therefore recorded identically for both the left- and right-bank of any section e.g. maximum/minimum river width and depth. Other variables such as maximum and minimum bank height and slope were entered separately for the left and right bank.

Bankside management

A record of the bankside management was included to give an idea of the land-use of the riverbank itself, irrespective of the surrounding land-use and other management included on the maps but beyond the bank proper. This classification was derived mainly from the schematic maps, taking into account the vegetation present on the bank, the presence of bankside fences etc. (to preclude/facilitate grazing) and any additional relevant information from the maps. It was thought appropriate to create this category as the management of the bank itself was often completely different to the land-use up to the bank's edge. This may have had some bearing on the distribution of the invasive species that would otherwise be masked. Four categories were entered to cover the various methods of riverbank management, these were:

- Unmanaged - all areas where no discernible management was noted, the bank generally being dominated with tall/rank vegetation or wood/scrub.
- Mown - all areas where mown grass was maintained to the water's edge.
- Meadow - all areas where hay/silage was gathered down to the water's edge
- Grazed - all areas that were grazed down to the water's edge, even if rank vegetation still persisted on some areas of the grazed bank.

Land-use, substrates and riverbank features

Surrounding land-use was recorded both in the descriptive section of the RCS surveys and also on the schematic maps. Both of these sources of information were consulted and the appropriate land-use categories from those listed in Appendix I (Table AI.4) recorded in terms of presence/absence.

Various substrate types were recorded (Appendix I Table AI.5) indicating the range of deposition and erosion environments in a section and the constituent substrates of the riverbank.

All riverbank features incorporated into the schematic maps (see Appendix I, Table AI.3) were recorded on a scale ranging from 1-10 when present, each increment in the scale representing an increase in occurrence along the bankside of ten percent.

It was important to make a distinction between the two categories of woodlands and isolated trees when interpreting the maps. The woodland category was recorded for any lengths of the bank where there was a continuous cover of trees depicted along the edge of the river. When the trees were depicted as individual trees or as isolated pairs along the bank they were classified as isolated trees. It was assumed that areas of woods would create different light and sub-canopy conditions for plant life than would isolated trees.

Macrophyte data

The abundance scores for the macrophytes followed the DAFOR scale of abundance: 1 representing rare through to 5 for species which dominated the bank. Macrophyte data were mostly derived from the RCS macrophyte summary page along with any additional macrophyte records from other parts of a section survey. The lack of non-arboreal plant recording in general on some stretches resulted in a somewhat depauperate data-set for several species.

3.2.2 Data reduction

Because of the large number of variables in the RCS data some reduction and amalgamation of variables was necessary. Variables that occurred in less than ten percent of all the sections of any catchment were either omitted from subsequent analysis or amalgamated into composite variables.

Newly created variables were designed so as to amalgamate similar variables, which served both to increase the number of occurrences of variables (i.e. lumping several similar variables with very low occurrences) and also to create more ecologically meaningful variables. For example, if a species was very dependent upon disturbance to exist in a habitat there are several variables that could create suitable disturbance. Individually they may not occur in sufficient quantity to indicate a significant statistical correlation with this species but if all such variables are amalgamated, the resultant composite variable may prove a good predictor.

The composite variables shown in Table 3.3 were created with the aim of amalgamating variables that may have similar ecological effects upon a study species.

Table 3.3 The composite variables created to amalgamate related groups of variables (see Appendix I, Table AI.7 for details). These variables were recorded in terms of presence/absence only.

Variable	Description
Urban	contains all urban variables
Disturbance	contains all of the areas of disturbance or past disturbance along the river.
Woodland	contains all the types of woodland encountered along the river.
Mixed Woodland	contains all woodland with conifers.
Farming	incorporates the various types of farmland found along the river.
Wilderness	A variable including all of the wild or semi-natural habitats (other than woodland) found along the length of the river.
Amenity/Mown/ Gardens	contains variables characteristic of managed landscapes.
Road/Bridge	combines these two access variables
Hard Cliffs	amalgamates both hard earth and rock cliffs
Herb/Ruderal	amalgamates sparse herb/ruderal and herb/ruderal vegetation
Grass/Herb	amalgamates sparse grass/herb and grass/herb vegatation

The occurrence of the study species on the bankside and within the river itself were combined for each section.

Accurate distribution information for the study species was collected over the summer of 1996 for the Wear catchment (see Chapter 5 for detail), adding previously over-looked populations and any expansion that had occurred since the initial surveys. These data were combined with the RCS data as a means of producing and testing models of species occurrences.

3.2.3 Analysis techniques

The distributions of the study species on the two catchments were mapped using DMAP (Morton 1998). For the Wear catchment the changes in distribution of the species between the RCS survey and the current survey were compared.

After reduction and amalgamation of the RCS variables, the resultant variables were examined. For quantitative variables the degree of skew and kurtosis was used to assess normality, transformations being undertaken on non-normally distributed data. As no

variables were successfully transformed to correspond to a normal-type distribution Mann-Whitney U-tests were used to determine which variables were related to the study species' distributions. Similarly χ^2 tests of heterogeneity were undertaken on the binary and categorical variables to see how they related to species' distributions.

Spearman-rank correlation analyses were used to highlight correlations between the RCS variables, as the occurrence of very strongly correlated variables could lead to some variables becoming redundant in further analyses and may prove useful in subsequent interpretation of variables selected in predictive modelling.

The distributions of the three species in relation to subsets of RCS variables were examined using χ^2 analysis, the subsets being:

i. Morphological features
ii. Bankside features
iii. Surrounding land-use
iv. Bank management
v. Composite variables.

3.3 Results

3.3.1 Catchment characterisations derived from RCS variables

A selection of bankside variables derived from the RCS data, along with altitude, are included in Figure 3.4, splitting the data into those of the upper, middle and lower areas of each catchment, and also differentiating tributaries of the two main rivers. These indicate that although the two catchments are very similar in character there are some differences between them. The tributaries of the two catchments are also very similar in character.

The upper reaches of the Tees catchment have a higher mean elevation than do those of the Wear. This is a result of the Tees RCS survey continuing higher up the catchment than that of the Wear and does not reflect a higher origin of the Tees. The width profiles along the catchments highlight that fact that the Tees is the larger of the two rivers. The Tees is consistently more wooded than the Wear, though the Wear has a greater proportion of coniferous plantation along its banks. The lower reaches of the Tees have more herb/ruderal vegetation than do those of the Wear; the latter having more grass/herb vegetation on average along its entire length. Both rivers have large areas of artificially created banks in the lower reaches, though increased densities of bridges along the lower Wear suggest a more urban catchment. The increased grazing of bankside along the Tees may have implications for alien species spread if species are grazing intolerant, though similar amounts of unmanaged bankside along both catchments suggests alternative habitat are equally available in both catchments.

The summaries of surrounding land-use (Figure 3.5) show that both catchments contain a similar amount of woodland, though the lowlands of the Wear are more wooded than those of the Tees. The increased occurrence of arable land adjacent to the Tees reflects more intensively farmed lowlands. Although the Tees has more urban areas adjacent to the river in the lowlands, all the other anthropogenic characteristics (roads, disturbance and areas of dereliction) are more frequent along the Wear catchment.

Figure 3.6, showing substrate distributions along the catchments, reveal the Tees catchment to harbour more silt in its lower reaches and along its tributaries than the Wear; the latter tending to have a greater proportion of mud and sand than the Tees in all areas of the river. The Tees tends to have more shingle areas than the Wear, especially in the middle reaches, though there is no obvious difference in the occurrence of larger substrates (not illustrated) between the two catchments.

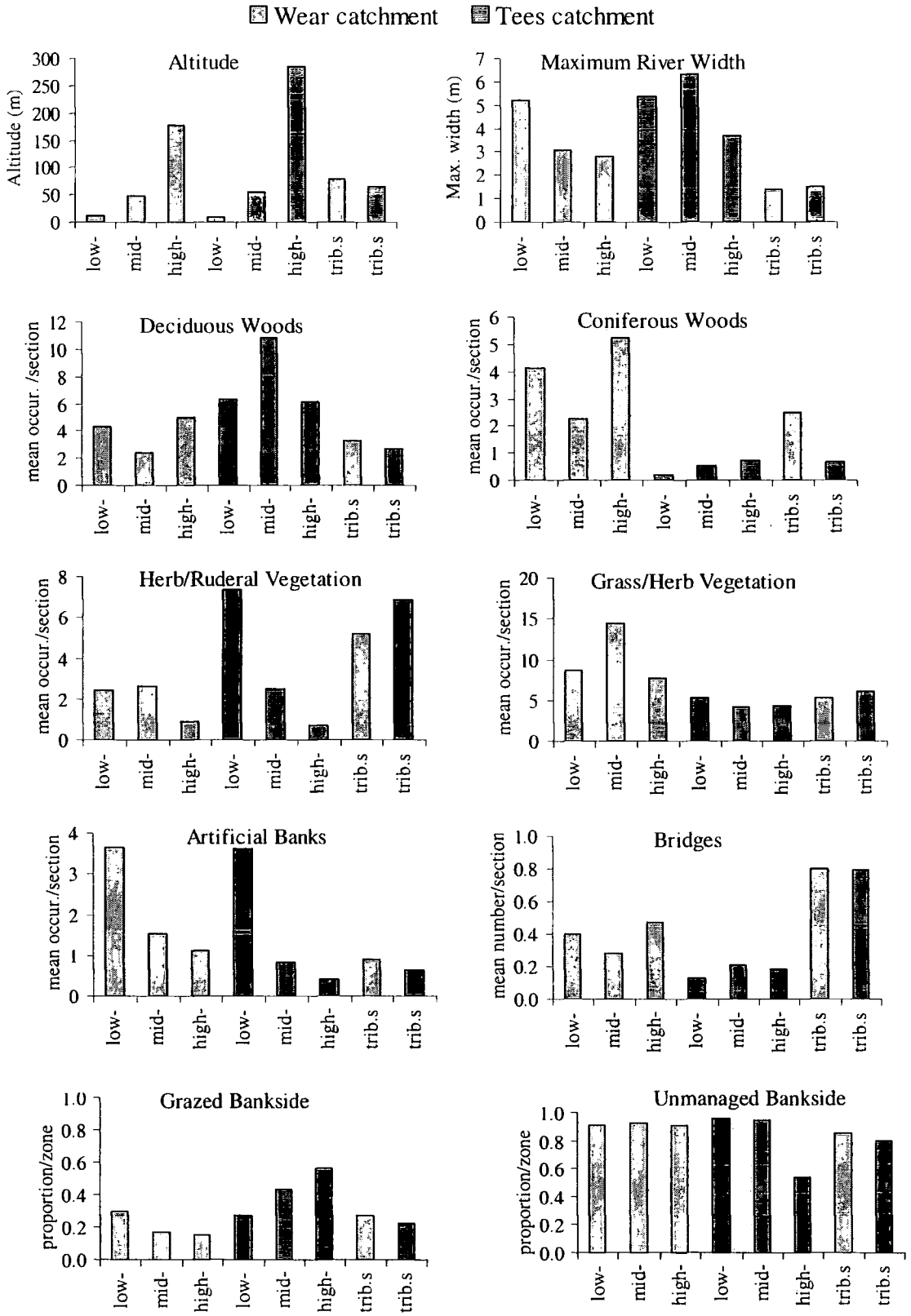

Figure 3.4 Differences in the bankside characteristics of the Tees and Wear catchments. Mean values are shown for sites in the lower, mid and top third of each of the main rivers along with mean values for the RCS tributaries in each catchment. Altitude and river width use continuous data. Values for deciduous & coniferous woodland, herb/ ruderal vegetation, grass/herb vegetation and artificial banks all refer to occurrences in 50m sections within a 500m RCS section (i.e. values in relation to a potential maximum occurrence of 20). Grazed and unmanaged bankside were recorded only in terms of presence/absence, values therefore represent proportional occurrence in RCS sections.

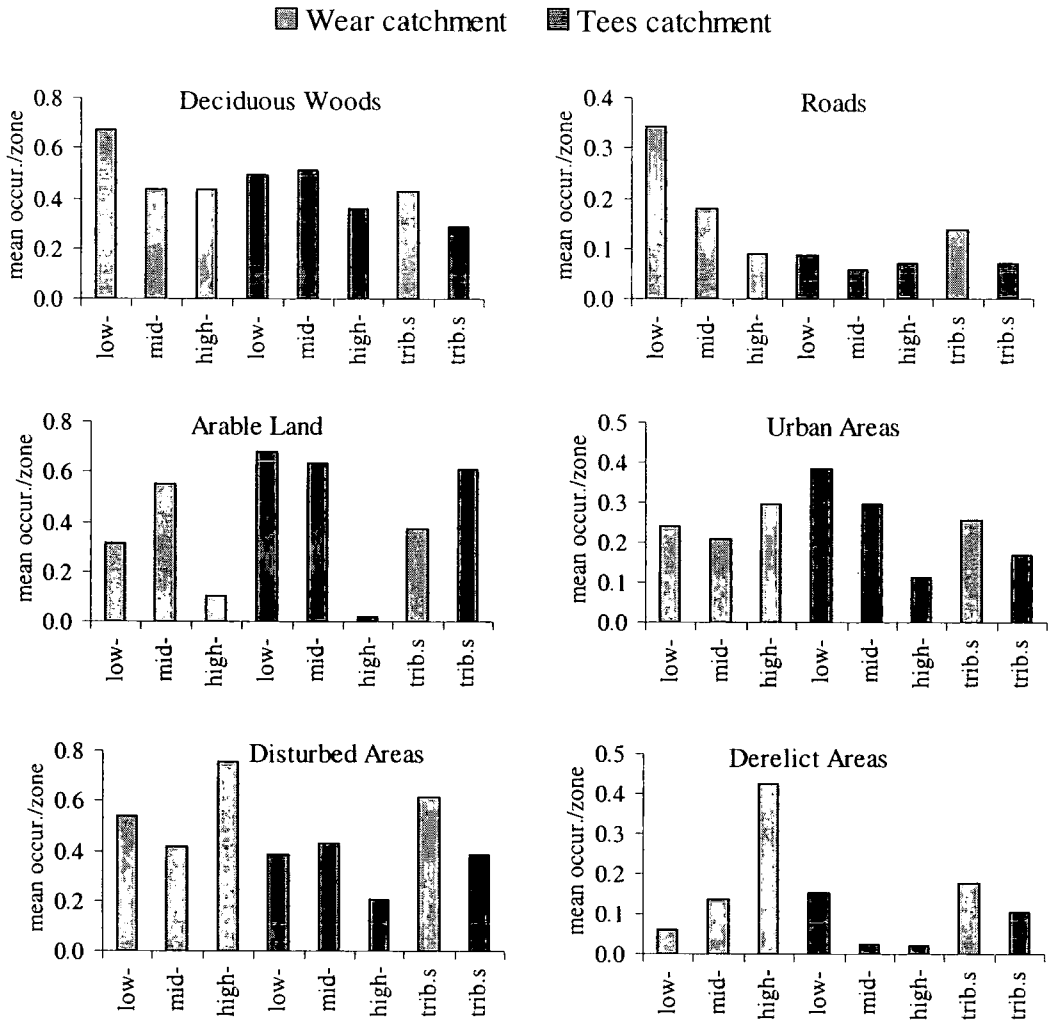

Figure 3.5 Differences in the surrounding land-use characteristics of the Tees and Wear catchments. Mean values are shown for sites in the lower, mid and top third of each of the main rivers along with mean values for the RCS tributaries in each catchment. All the variable are recorded in terms of presence/absence only, values therefore represent proportional occurrence in RCS sections

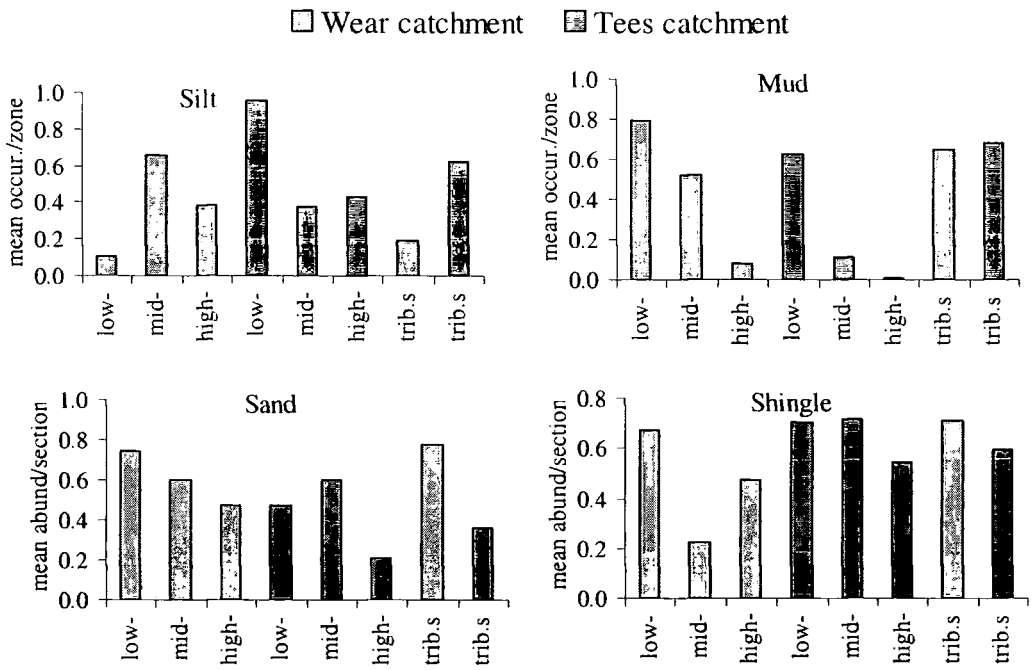

Figure 3.6 Differences in the riverine substrates of the Tees and Wear catchments from the RCS data. Mean values are shown for sites in the lower, mid and top third of each of the main rivers along with mean values for the RCS tributaries in each catchment. All the variable are recorded in terms of presence/absence only, values therefore represent proportional occurrence in RCS sections

3.3.2 Species distributions

The distribution patterns of *Impatiens* along the two catchments (Figure 3.7. & 3.8) show several similarities. In both catchments the species avoids the inter-tidal lowland reaches of the main rivers and also the reaches at higher altitudes. Within these limits the species appears to be continuously present from the downstream limit of the main rivers for some distance upriver, after which its occurrence becomes more sporadic. The altitudes relating to the start of its more disjunct distribution are similar for both catchments, about 140 m a.s.l. on the Tees and 160 m a.s.l. on the Wear. Its absolute upper-limit is also remarkably consistent in both catchments, at about 230m a.s.l. Distribution of *Impatiens* along the smaller tributaries is less consistent between catchments. Along the Wear most smaller tributaries within its altitudinal bounds are fully colonised, with perhaps the exception of the upper Gaunless (see Figure 3.1 for river locations). Along the Tees however the pattern is not so clear. Some small tributaries in an apparently suitable climatic zone remain completely uncolonised whilst others are only partly so. The Skerne for example is only partly colonised and the species is completely absent from Clow Beck. Absence of the species in the tributaries

around Teesmouth may be indicative of inter-tidal conditions, though its occurrence in the lowest reach of some such tributaries contradicts this.

The sites of *Heracleum* on both catchments also show similarities (Figures 3.9 & 3.10). Like *Impatiens*, *Heracleum* also avoids the inter-tidal zones of the rivers. Also like *Impatiens*, *Heracleum* has a zone of almost continual occurrence along both main rivers and an upstream zone of only sporadic occurrence. On both rivers the upstream limit of the continual distribution occurs at about 30m a.s.l. whilst the absolute upper limit is 40m a.s.l. on the Wear and 80m a.s.l. on the Tees. Along both catchments the species occurs almost exclusively on the main river and avoids the smaller tributaries.

Compared to *Heracleum* and *Impatiens* the distribution of *Fallopia* is much more patchy along the study catchments (Figures 3.11 & 3.12). Indeed along the Tees it is only rarely recorded. However this may not now be the case, as the recent River Wear surveying indicated a great increase in species range compared to the RCS survey. It is worth noting that most of the new records of *Fallopia* along the Wear occurred downstream of originally recorded sites, perhaps suggesting a real downstream expansion due to vegetative fragmentation. The observed small size of many of these new colonies reinforces this idea.

The altitudinal limits of *Fallopia* along the two catchments are vastly different. Along the Tees it only occurs to 30m a.s.l., whereas on the Wear it occurs sporadically up to 170m a.s.l., with an isolated record at 220m a.s.l.

The differences in sites occupied by the three species on the Wear between the original RCS and the present survey show apparent range expansion for all three species, with no loss of sites between surveys (see Discussion for apparent *Fallopia* losses).

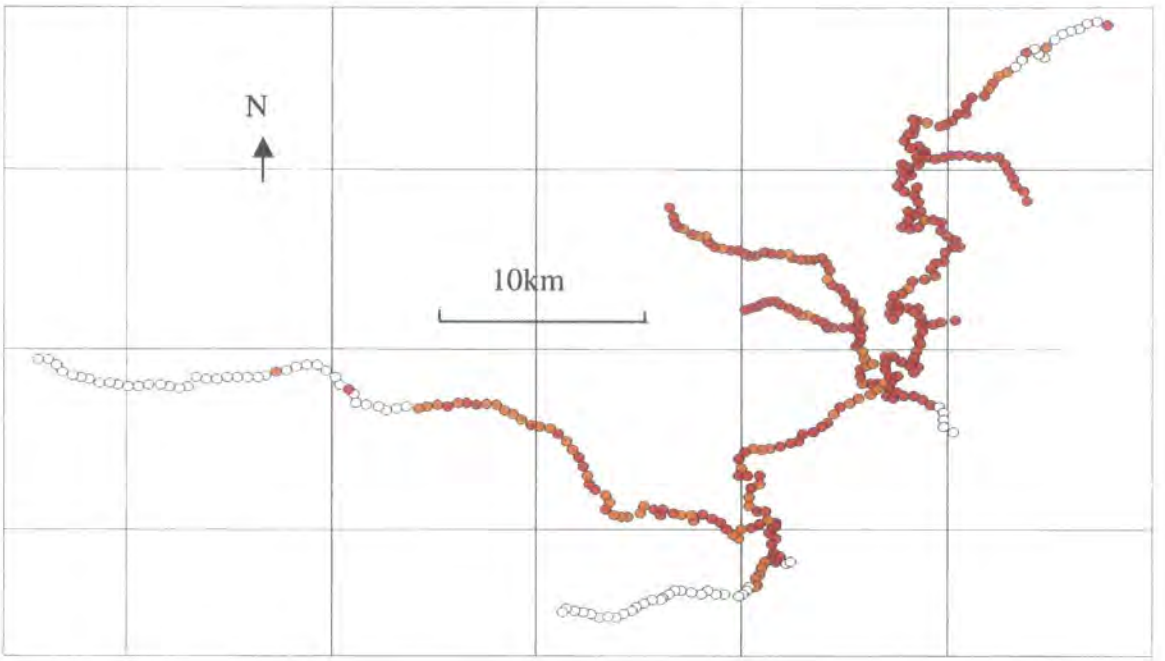

Figure 3.7 Distribution of *Impatiens* along the Wear catchment derived from both RCS data and recent surveys. Original RCS sites are coloured red, additions from the recent surveys are coloured orange.

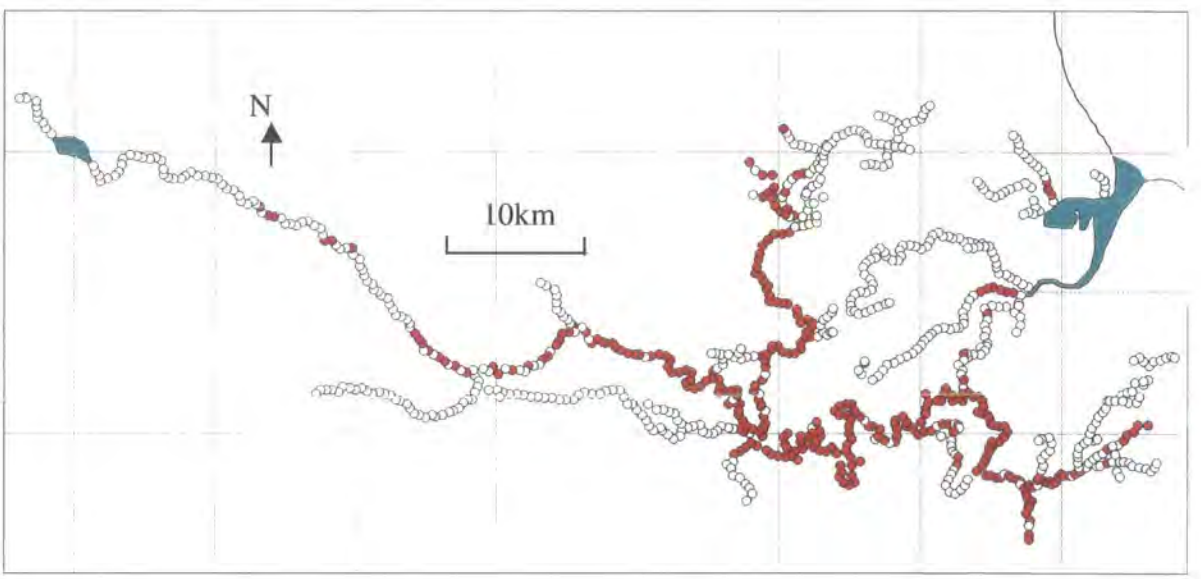

Figure 3.8 Distribution of *Impatiens* along the Tees catchment derived from RCS data only. Sites of recorded occurrence are coloured red.

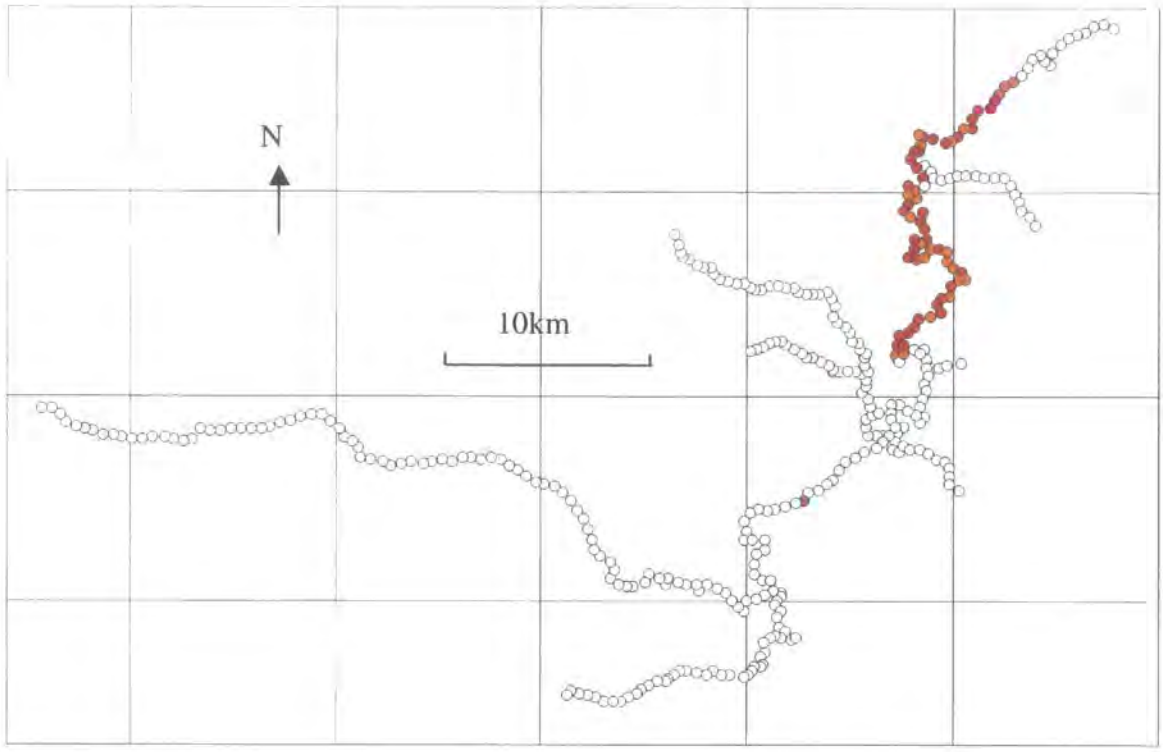

Figure 3.9 Distribution of *Heracleum* along the Wear catchment derived from both RCS data and recent surveys. Original RCS sites are coloured red, additions from the recent surveys are coloured orange.

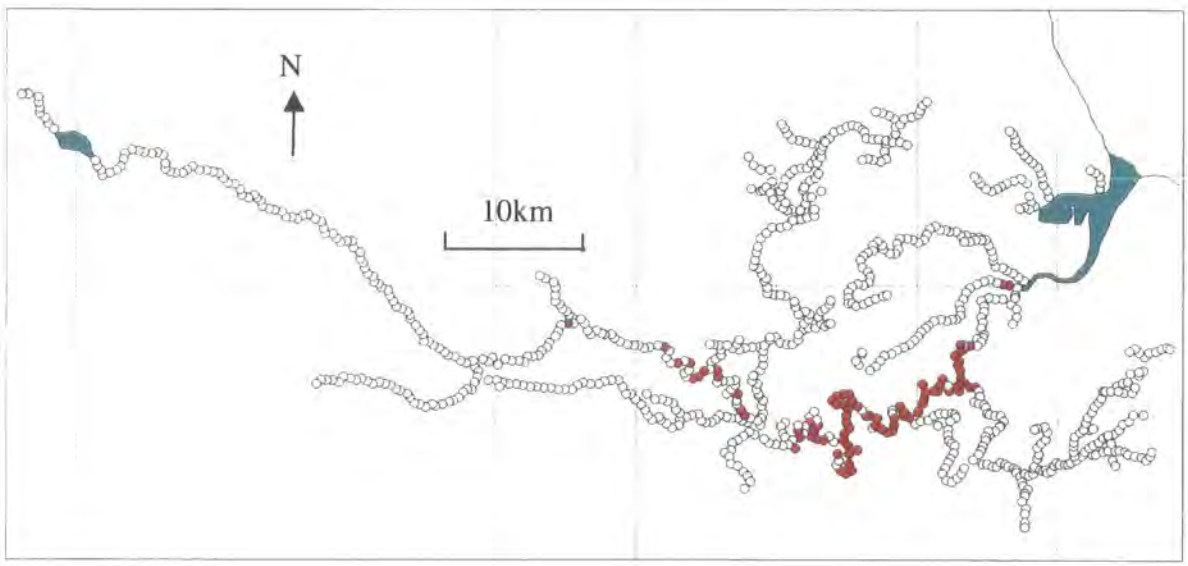

Figure 3.10 Distribution of *Heracleum* along the Tees catchment derived from RCS data only. Sites of recorded occurrence are coloured red.

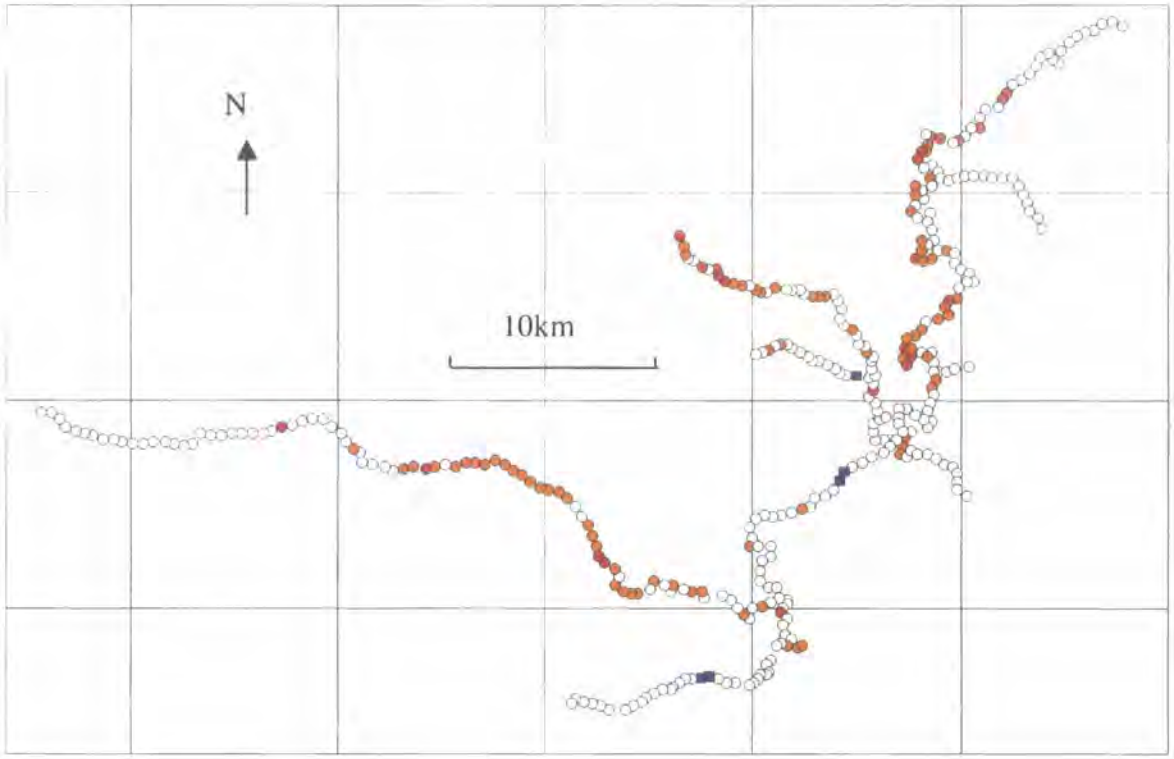

Figure 3.11 Distribution of *Fallopia* along the Wear catchment derived from both RCS data and recent surveys. Red circles represent original RCS sites that also contained *Fallopia* during the recent survey. Blue squares represent sites from the original RCS data but from which the species is now absent (but see text). Orange circles represent additions from the recent surveys.

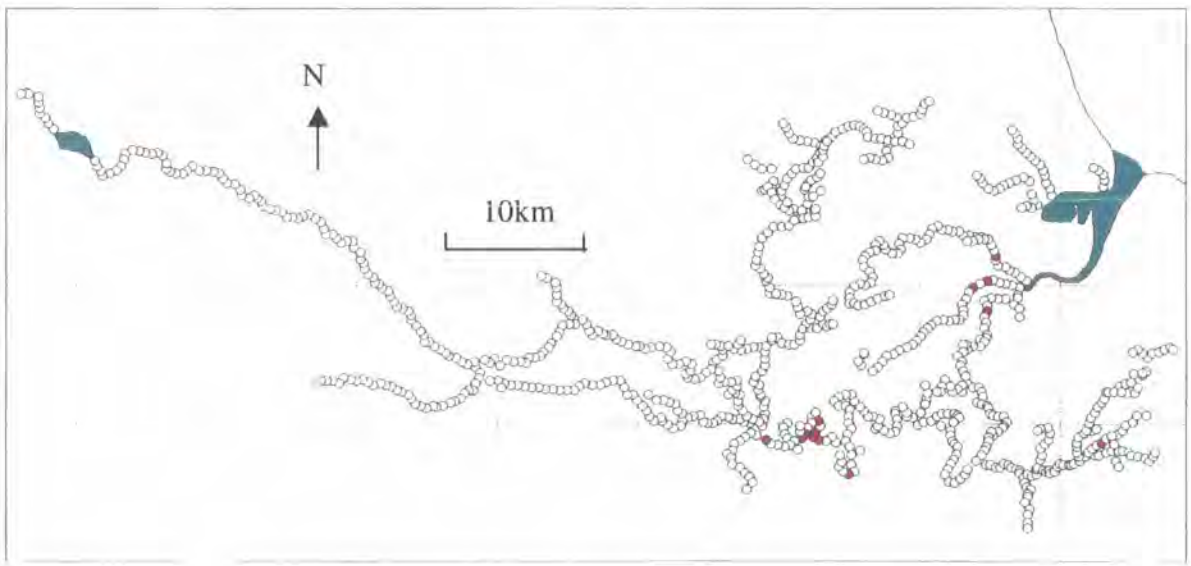

Figure 3.12 Distribution of *Fallopia* along the Tees catchment derived from RCS data only. Sites of recorded occurrence are coloured red.

3.3.3 Summary statistic

The results of the exploratory χ^2 and Mann-Whitney U-tests of the RCS variables on the Wear and Tees catchments with the species distributions are included in Tables AII1-AII4 in Appendix II. The major trends of these tables however are summarised in Table 3.4 below.

Table 3.4 Variables from the RCS found to have either no effect on a species in both catchments (No effect) or to significantly affect a species' distribution on both catchments (Significant).

a) *Impatiens*

No effect	Significant	
Conif_ba	Alt.	***
Conif_lu	Gr/he	***
Derelict	He/ru	***
Grazed	Landuse	***
Hard_cls	Max_dep	***
Max_slop	Min-dep	***
Mixed	Silt	***
Reeds	Soft/slum	***
Roads	Unman.	***
Scrub_ba	Tot_wood	***
Scrub_lu	Artific.	**
Weirs	Decid_ba	**
Wild	Decid_lu	**
	Max_ht	**
	Arable	*
	Islands	*
	Min_ht	*

b) *Fallopia*

No effect	Significant	
Gr/He	Herac.	**
Am/Mo/Ga	Impat.	**
Arable	Max_slop	**
Bridge	Farming	*
Cobb_tot		
Conif_ba		
Conifer_lu		
Dep_max		
Derelict		
Disturb		
Hard_cls		
Islands		
Landuse		
Min_ht		
Min_Slop		
Mixed		
Mud		
Reed		
RoadorBr		
Sand		
Scrub_ba		
Scrub_lu		
Sh_mo		
Silt		
Stream		
Substrate		
Weir		
Wild		

c) *Heracleum*

No effect	Significant	
Am/Mo/Ga	Alt.	***
Artific.	Boul	***
Bedrock	Decid_lu	***
Conif_lu	Impat.	***
Graze	Landuse	***
Scrub_lu	Max_wid	***
Stream	Min_dep	***
Tot_res	Min_wid	***
	Tot_wood	***
	Decid_ba	**
	Disturb	**
	Fallop.	**
	Hard_cls	**
	Soft/sl.	**
	Farm	*
	Min_slop	*
	Mixed	*
	Reeds	*
	Silt	*
	Substrate	*

Variables are as described in Appendix I, Table AI1. The minimum value of significance of any variable which influenced a species on both catchments is indicated by *, ** and *** for 0.05, 0.01 and 0.001 levels respectively

The summary statistics indicate that there are significant associations of the study species with a whole variety of the RCS variables including bankside substrates, bankside habitats and surrounding land-use, as well as other variables such as altitude and physical bankside parameters. It therefore seems that no one set of variables is especially related to the distribution of the three species. The wide range of variables significant on both catchments, at least for *Impatiens* and *Heracleum*, suggests that

several factors will be necessary to predict species occurrences. The large set of non-significant variables for *Fallopia* may be indicative of a species with more catholic habitat preferences or simply be a result of a currently limited distribution. Unlike the other two species neither bankside substrate nor surrounding land-use are found to influence the occurrence of *Fallopia*. Similarly many of the variables reflecting anthropogenic influences are deemed unimportant.

For both *Impatiens* and *Heracleum* altitude is one of the most important variables influencing distribution. The lack of importance of this variable for *Fallopia* reflects the more disjunct distribution of the species. Many of the variables most associated with the occurrence of *Impatiens* reflect habitat preferences, hence the inclusion of various woodland habitats, grass/herb and herb/ruderal habitats, and unmanaged banks. Others such as land-use and river depth are more likely to be indicative of a climatic zone along the catchment.

The variables that significantly affect *Heracleum*'s distribution are similar to those for *Impatiens*, hence its inclusion as a significant variable in *Impatiens*' distribution. Habitat preferences of *Heracleum* are reflected by the inclusion of woodland categories, substrate type and disturbed and soft/slumping banks as significant variables. Conversely the inclusion of boulders and rock/hard earth cliffs may indicate avoided habitats. The inclusion of substrate types could indicate either habitat selection, deposition environments or may simply reflect the stage of a river and hence reflect a climatic zone.

Of the few variables that influence the distribution of *Fallopia* along the two catchments two of the most important are the other invasive species. There is a suggestion that all of these invasive species rely on certain riparian characteristics (perhaps regularly flooded disturbed banks), the species all prospering where such conditions arise. Further differences in distribution of the three species may be due to secondary factors such as their relative climatic tolerances and differing competitive and dispersal abilities.

The RCS variables from both catchment databases were examined for correlations, both between each other and also with the alien species. This could lead to some variables becoming redundant and indicate if any variables shown to influence a species occurrence (from above) act in opposite directions on the two catchments. Additionally the detection of correlations may be of assistance in the interpretation of analyses when apparently spurious variables are indicated to be important.

Correlations among the entire suite of variables were calculated using Spearman-rank analyses. The results of these correlations are given in Appendix II (Tables AII5 & AII6) and indicate a large degree of correlation between variables in both catchments. This could lead to problems in the interpretation of results to determine driving variables in a species' distribution. It is therefore logical to remove one of any pair of strongly correlated variables from the variable lists for further analysis, thereby simplifying calculations and interpretation.

Variables were selected as being strongly correlated when they had a correlation co-efficient greater than the value of 0.4. This value was selected as the best compromise to exclude highly correlated variable pairs without removing a large proportion of the variables (most variables within the riparian system were correlated to some degree). Variables falling into this category in both catchments (excluding situations when a variable is derived partly from another i.e. amalgamated variables) are as follows:

Maximum and minimum river width,
Maximum and minimum river depth,
Maximum and minimum bank height,
Boulders and cobbles,
Heracleum and max/min river width.

One of each of these pairs of correlated variables could be legitimately removed from further analysis. For the first four pairs, the variable that was correlated with the least number of other variables was retained. If this was similar for the two variables then the most ecologically sensible variable was retained. The variables dropped from the analysis at this stage were minimum river-width, maximum river-depth, maximum bank-height and boulders. Other pairs of variables which were correlated in both catchments but not to the same degree (one of the co-efficients is <0.4) are:

Heracleum and Altitude,
Impatiens and Altitude,
Mud and Altitude,
Bedrock and Hard cliffs,
Boulders and Deciduous wood,
Boulders and Mud,
Cobbles and River width,
Conifer (ba) and Deciduous (ba),
Conifer (ba) and Deciduous (lu),
Deciduous (ba) and Deciduous (lu),
Deciduous (ba) and Maximum river width,
Maximum river depth and River width,
Minimum river depth and River width,
Grazed bank and Unmanaged bank,
Grazed bank and Grassland,

Grazed bank and short /mown bank,
Maximum bank height and Maximum river width,
Minimum bank height and Minimum river width.

The first two pairs of correlations reflect the preference of these two study species for the lower reaches of the catchment, as seen earlier. Some of the other weaker correlations are features of the nature of the catchments i.e. as the river gets wider, it also gets deeper and the banks get taller. Other pairs of variables are also logical e.g the relationship between grazed banks and grassland, and grazed bank and short grass. Other correlations are related to the zones of a river e.g. the relationships of decreasing mud with increasing altitude, the negative correlation between boulders and mud etc.

There are correlations between the three alien species although these are not as strong as the correlations mentioned above. This may again indicate a similarity in habitat choice.

Heracleum and *Impatiens* are positively correlated with many variables which are typical of the lower regions of the two catchments e.g. river width, river depth, ruderal vegetation, arable land, deciduous woodland and bank height. Conversely they are negatively correlated with features of the upper catchment e.g. increasing altitude and grasslands. Some of these variables may reflect habitat selection rather than just the particular zone of the river e.g. deciduous woodland, ruderal vegetation.

The majority of the other weaker correlations between variables can be ascribed to the differences in land-use and climate with increasing altitude along the catchments.

3.3.4 Management regime

A comparison was made of the ability of grazed and unmanaged riverbank areas to support the three species in both catchments. For each species χ^2-tests examined the occurrence of species in relation to management on each bank to see if the observed occurrences differ from an expected distribution.

The distribution of *Fallopia* is not significantly different from expected on either catchment (Wear χ^2: n=678, df=2, P>0.05. Tees χ^2: n=1543, df=2, P>0.05). *Impatiens* occurs significantly less than expected in grazed areas on both catchments (Tees χ^2: n=1543, df=2, P<0.01. Wear χ^2: n=678, df=2, P<0.01), whereas *Heracleum* occurs significantly less than expected in grazed areas on the Tees catchment only (χ^2: n=1543, df=2, P<0.01).

The distribution of the two bank management categories indicates that unmanaged banksides are the norm for much of the two catchments, with grazed banks only occurring with any frequency in the upper catchments.

3.3.5 Morphological features

The distributions of the species in relation to bankside substrates along the two catchments were examined to see if any species preferences were apparent. The figures presented (Figure 3.13 & 3.14) display the number of sites in which a species occurs in association with different substrates. Also plotted is the number of such sites a species would be expected to occur in given no substrate selection i.e. assuming proportional distribution in relation to substrate frequency.

Heracleum showed habitat preferences in both catchments, though only on the Wear did this differ significantly from expected (Figure 3.13). On both catchments it avoids areas of boulders and bedrock, whilst there is an apparent difference in its preference of smaller substrates between the two catchments. On the Wear it occurs more than expected in areas of mud and less than expected in silty sections. However on the Tees the species occurs much less than expected in areas of mud. These latter apparent conflicts may arise due to differences in recording substrates between the two catchments or may reflect the greater occurrence of unoccupied tributaries around the Tees estuary (primarily lined with mud). Similarly *Impatiens* on the Tees occurs less than expected in areas of mud and silt (Figure 3.14), a trend not apparent on the Wear. All three species occur less than expected in areas of bedrock on both catchments.

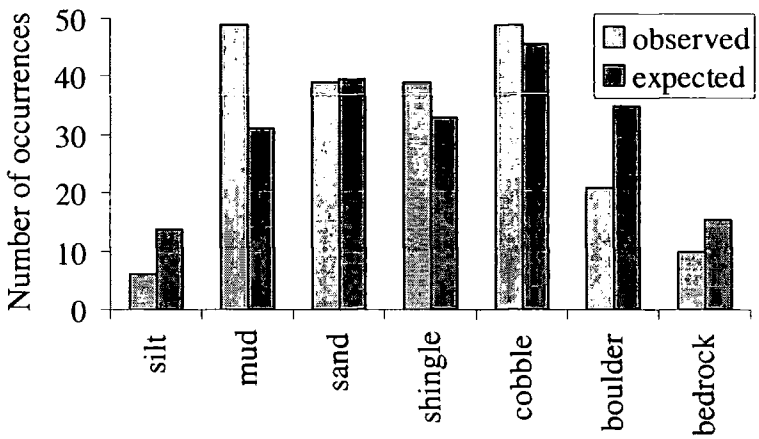

Figure 3.13 The number of sites of occurrence of *Heracleum* on the Wear catchment in relation to morphological features. Also shown are the expected occurrences given no preference (χ^2: n= 1329, df= 6, P<0.01).

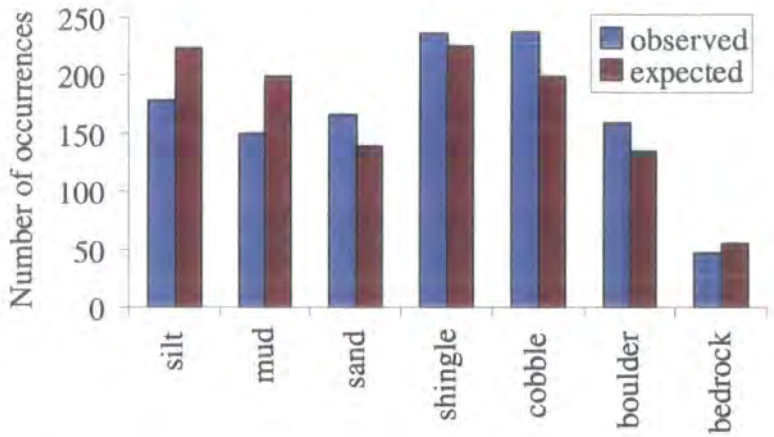

Figure 3.14 The occurrence of *Impatiens* on the Tees catchment in relation to morphological features. Also shown are the expected occurrences given no preference (χ^2: n= 2781, df= 6, P<0.01).

Fallopia showed no significant difference from an expected distribution in relation to bankside substrates on either catchment.

3.3.6 Bankside features

The occurrence of *Heracleum* in relation to riverbank habitats on both catchments shows significant deviations from expected (Figure 3.15), calculated using the same method as in section 3.3.4. A slight preference for areas of woodland and herb/ruderal vegetation and an avoidance of grass/herb vegetation are factors common to both catchments.

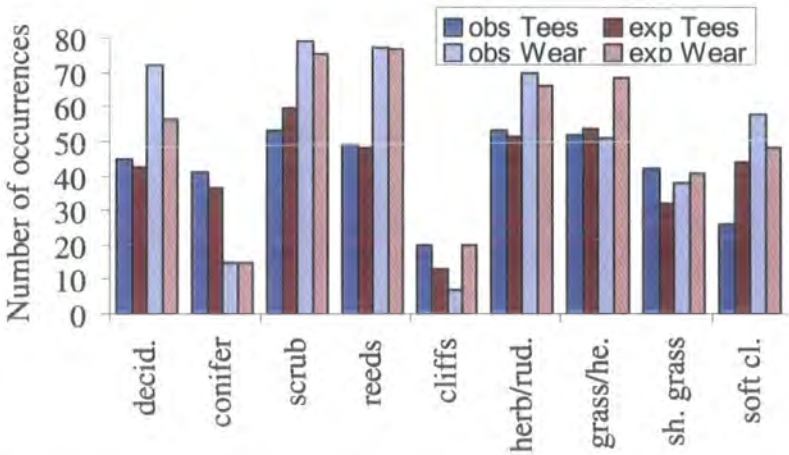

Figure 3.15 The occurrence of *Heracleum* on the Wear and Tees catchments in relation to bankside features. Also shown are the expected occurrences given no preference (Wear χ^2: n= 2006, df= 8, P<0.05. Tees χ^2: n= 3483, df= 8, P<0.05).

The occurrence of *Impatiens* in relation to bankside habitats is only significantly different from expected on the Tees (Figure 3.16), occurring more than expected in

areas of woodland and herb/ruderal vegetation and less than expected in areas of grass/herb vegetation. A preference for herb/ruderal vegetation was also apparent on the Wear catchment.

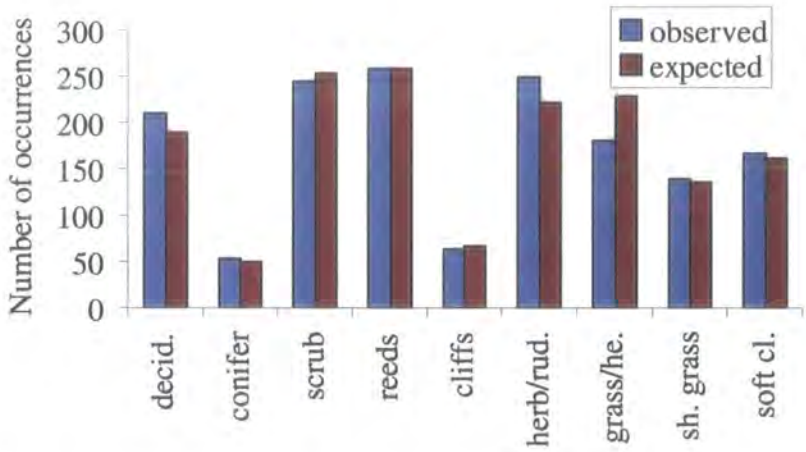

Figure 3.16 The occurrence of *Impatiens* on the Tees catchment in relation to bankside features. Also shown are the expected occurrences given no preference (χ^2: n= 3483, df= 8, P<0.05).

Fallopia showed no significant difference from an expected distribution in relation to bankside habitats on either catchment.

3.3.7 Surrounding land-use

The distribution of *Heracleum* in relation to surrounding land-use (Figure 3.17) shows significant variation from that expected in both catchments. It occurs more than expected in areas by deciduous woodland and less than expected in areas by grasslands and adjacent to roads/bridges.

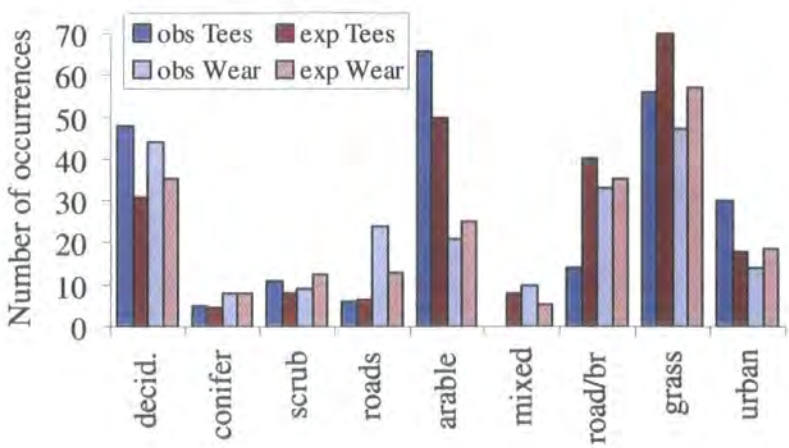

Figure 3.17 The occurrence of *Heracleum* on the Wear and Tees catchments in relation to surrounding land-use. Also shown are the expected occurrences given no preference (Wear χ^2: n= 944, df= 8, P<0.01. Tees χ^2: n= 2268, df= 8, P<0.01).

The distribution of *Impatiens* with regards surrounding land-use is significantly different from expected only on the Tees catchment (Figure 3.18). Here it occurred more than expected in areas adjacent to deciduous woodland and urban areas and less than expected in areas by grasslands and roads/bridges. Along the Wear it occurs more than expected by areas of deciduous woodland and arable areas but less than expected by urban areas.

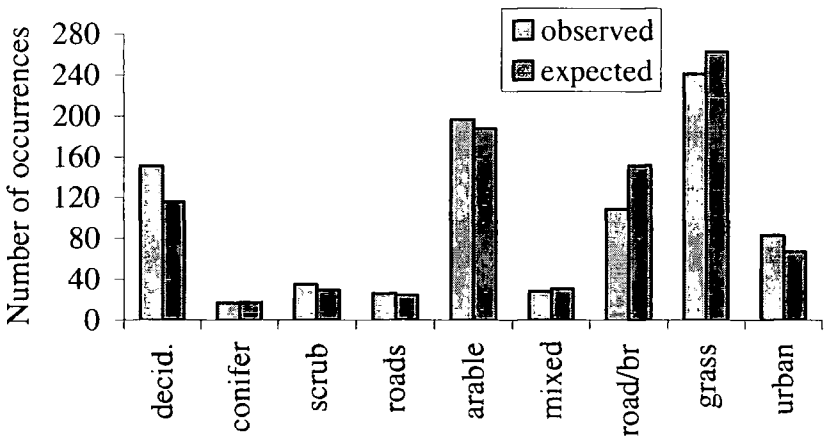

Figure 3.18 The occurrence of *Impatiens* on the Tees catchment in relation to surrounding land-use. Also shown are the expected occurrences given no preference (χ^2: n= 2268, df= 8, P<0.01).

Fallopia only deviates from its expected distribution in relation to surrounding land-use along the Tees catchment (Figure 3.19). Here it occurs more than expected by areas of deciduous woodland and adjacent to urban areas and roads; occurring less than expected by grassland areas.

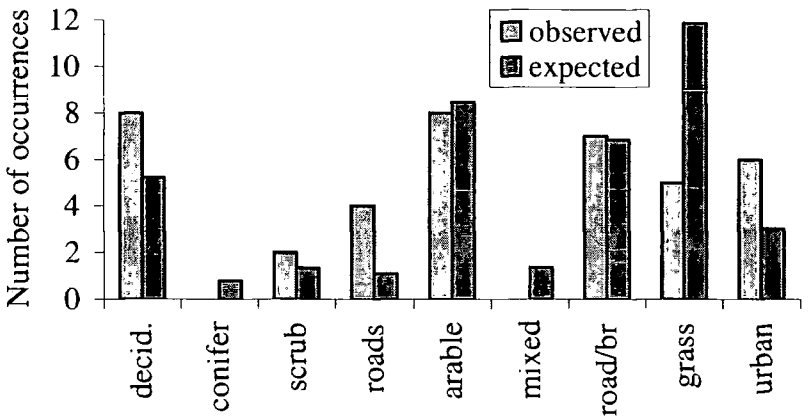

Figure 3.19 The occurrence of *Fallopia* on the Tees catchment in relation to surrounding land-use. Also shown are the expected occurrences given no preference (χ^2: n= 2268, df= 8, P<0.05).

3.3.8 Composite variables

The occurrence of *Heracleum* in relation to the composite variables is significantly different from expected only on the Tees catchment (Figure 3.20). There it occurs more than expected in areas of woodlands, amenity use and residential areas. Conversely it occurs less than expected in areas of disturbance, human activity and derelict and wilderness areas. Along the Wear it also occurs less than expected in areas of dereliction.

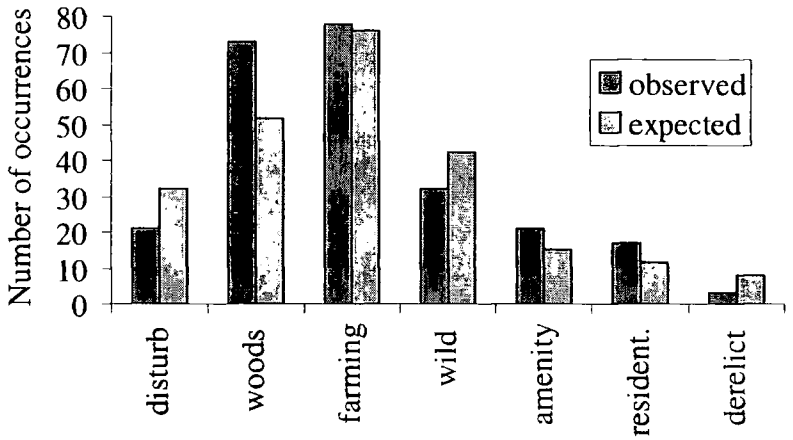

Figure 3.20 The occurrence of *Heracleum* on the Tees catchment in relation to composite variables. Also shown are the expected occurrences given no preference (χ^2: n= 2856, df= 7, P<0.01).

The occurrence of *Impatiens* in relation to the amalgamated variables is, like *Heracleum*, only significantly different from expected on the Tees (Figure 3.21). Here the species is seen more than expected in areas of woodland and amenity and residential areas but less than expected in wilderness and derelict areas. The trend of occurring more than expected in wooded areas is mirrored on the Wear catchment.

The distribution of *Fallopia* with regards the amalgamated variables does not significantly differ from expected on either catchment though on the Tees it occurred less than would be expected in farmland and greater than expected in amenity land-use and residential areas.

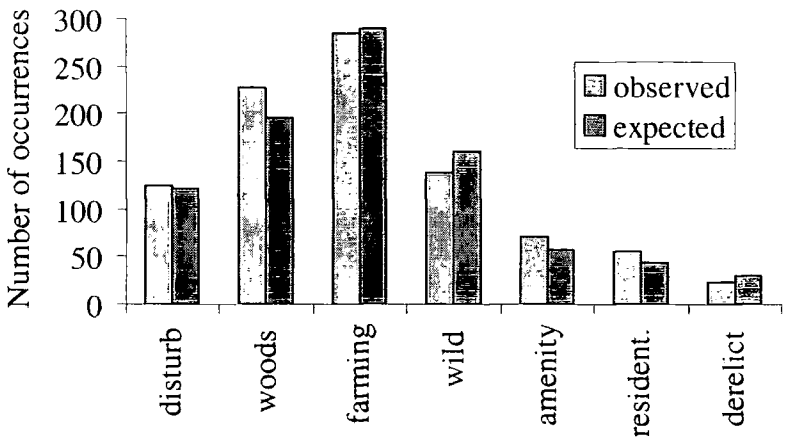

Figure 3.21 The occurrence of *Impatiens* on the Tees catchment in relation to composite variables. Also shown are the expected occurrences given no preference (χ^2: n= 2856, df= 7, P<0.01).

3.4 Discussion

One of the most important findings of this initial entry and analysis of data, and one which requires highlighting, along with the rider that all subsequent analyses are to some extent tempered by it, is that the RCS data are highly variable in quality. As such they are not ideally suited to the extraction of meaningful ecological data with which to examine species distribution. Two main flaws in the data are apparent. At the design level the selected habitat recording criteria are of limited use, both in terms of this work but also for their intended use as a reference tool. What for example is the difference between grass/herb and herb/ruderal vegetation? Nowhere in the recording methodologies are such terminologies explained. The level of ecological detail recorded by the RCS relies heavily upon additional information noted by a surveyor, which introduces the second flaw in the work; namely a lack of standardisation in recording detail. Just how useful data usually derived from a single visit are is also questionable. With standard recording schemes such as the butterfly monitoring scheme (Pollard 1979), and common bird census (Marchant 1983) requiring regular visits through a season, animal records included in the RCS data must be considered of little use. Even arboreal recording proved hopelessly inadequate in places, species rich hazel-coppice being unrecorded along one section of bank; instead a few sycamore trees were noted. To use the RCS as a conservation evaluation and management tool would be irresponsible, as the contained data mostly do not reflect the situation on the ground.

It must also be highlighted that the RCS surveys do not cover all riparian sites for the species in a catchment. Smaller tributaries, streams and ditches are overlooked by the

surveys, all of which could harbour the three study species, especially *Impatiens*, which thrives in many narrow channels.

Differences in distribution patterns of the species recorded both within and between catchments may arise due to many stochastic factors, though rational explanations can account for much of the patterning. The variation in colonisation of the smaller tributaries by *Impatiens* may be attributable to several factors. Firstly the absence of the species from tributaries close to the Tees estuary could be due to inter-tidal effects but the occurrence of the species at the base of one or two such tributaries would suggest otherwise. Also it seems that the lowest reaches of these estuary-bordering tributaries are at the upper-edge of the tidal limit and are therefore likely to be only marginally influenced by its effects. Another possibility is that propagules of *Impatiens* simply have not been able to reach these relatively isolated areas (the ability of the seeds to survive in saline conditions is unknown). A third possibility, and one which may explain absences throughout these tributaries, is that along such smaller waterways there is not the unmanaged bankside of the main rivers. A lack of fencing may allow grazing to take place down to the waters edge in some areas, thereby eliminating the species from entire sections of riverbank.

The similarity in *Impatiens*' upper limits on both catchments suggests altitudinal constraints to further colonisation and not dispersal limitations. The highest altitudinal records along both catchments are greater than those recorded in the species' biological flora. Beerling & Peerins (1993) state that the species has not been observed above 210m a.s.l.

The distribution of *Heracleum* along the two catchments indicates a marked preference for the lowlands and the large main rivers. The avoidance of smaller tributaries may be due to dispersal limitations or a lack of suitable habitats. Other work comparing tributaries to main rivers have shown reduced species diversity (Nilsson *et al.* 1994), which was attributable to substrate diversity and discharge. A lowland preference is mentioned in the species' biological flora (Tiley *et al.* 1996) but the recorded upper-limits of the species along the study catchments are much less than observed elsewhere (up to 213m, Tiley *et al.* 1996). This suggests that unlike *Impatiens*, *Heracleum* may be dispersal limited along the study catchments. This would have important implications should any additional upstream introductions occur.

The distribution of *Fallopia* along the catchments indicates that it too is currently very much dispersal limited in the two catchments. The catholic habitat tastes of the species

defy habitat unsuitability as being responsible for the species limited distribution. It is possible that along the Wear the upper limit of the species is climatically determined as the altitude of the upper limit fits values observed elsewhere. Beerling *et al.* (1994) state that the species rarely grows above 200m, suggesting lower limits for Scotland and Ireland. Other altitudinal limits observed by Beerling *et al.* (1994) include 320m in South Wales, 270m in North Wales, 230m in Shropshire and 305m on Dartmoor.

The difference in sites noted along the Wear for all three species between the RCS surveys and the current survey are remarkable. It is not certain whether these differences are attributable to actual species range changes or to poor recording in the initial survey. For *Impatiens* one suspects that the species was overlooked in some sites where its occurrence was only sporadic, though the large in-filling of sites in the upper catchment may also be due to large between year variation in seed deposition and over-winter survival. For *Fallopia* and *Heracleum* the in-filling may be due to real spread or initial under-recording, as most new records are of only small clumps which could have easily been overlooked or could just as easily be new sites of colonisation. Whereas with *Heracleum* the new sites fill in gaps in the previous species range, with *Fallopia* the new sites show spread into areas distant from previously recorded colonies. The apparent spread of the latter species in the upper Wear catchment from original sources is generally greater than in the lowlands, perhaps suggestive of increased fragmentation due to faster river flow and more frequent flash floods.

The apparent rapid spread of *Fallopia* from original sources, if real, sends a strong message for the early control of such species. The control effort necessary at the time of the original RCS survey would have been minimal compared to the effort now required to eliminate it.

Fallopia is the only species that seems to have disappeared from some sites between surveys. However this is probably erroneous as on revisiting apparently previously occupied sites the areas marked as containing the species always contained phenotypically similar species. For example several clumps of *Fallopia* marked on the RCS maps proved to actually be *Cornus sanguinea.* Other recorders apparently confused coppiced species such as *Corylus avellana*, *Ulmus* spp. and *Tilia cordata* with *Fallopia* stands.

The distribution of the three species in relation to riverine substrates shows a distinct avoidance of areas of bedrock, possibly due to inclement growing conditions or maybe

simply because of predominant upland occurrence of bedrock areas. The avoidance by *Heracleum* of reaches with boulders may also reflect a distribution tied more closely to the lowlands than the other two species. The wide tolerance by *Heracleum* and *Impatiens* of substrate types agrees with previous reports (Beerling & Perrins 1993, Tiley & Philp 1994). Similarly the distribution of these two species in relation to riverbank habitats also gives an indication of reported habitat preferences (woodland and herb/ruderal vegetation e.g. Beerling & Perrins (1993), Tiley & Philp (1994)).

The distribution of the species in relation to surrounding land-use seems to reflect either correlations with riverbank variables, the ability of a land-use type to sustain populations of the species or the climatic distribution of land-use types. Hence *Heracleum* occurs more than expected by areas of deciduous woodland, which is correlated with bankside deciduous woodland, a preferred habitat, but less than expected in areas of grassland and road/bridges, both of which are prevalent in the uplands and may therefore reflect a climatic limit. The situation is similar for *Impatiens,* which additionally occurs more than expected in urban areas.

In the case of *Fallopia* the apparent preference for locations by deciduous woodland is likely to reflect either source populations or a correlation with bankside woodland habitat. Conversely urban areas and adjacent roads are likely to reflect sources of introduction.

The distribution of the species in relation to the occurrence of amalgamated variables show similar trends with species occurring more than expected in preferred habitats and less so where unsuitable land-use adjoins e.g. derelict, wilderness and human influenced areas for *Heracleum* and *Impatiens*

The general conformity of the occurrence of *Fallopia* with the expected distributions in relation to these variable groupings may either arise from the species' catholic habitat preferences, the relatively low number of occurrences or from being limited by factors other than habitat.

Overall it can be seen that the distribution of these introduced species is influenced to some extent by a wide variety of riverine variables. As a result of this it should be possible to use such variables to create models to predict their occurrence both in occupied and unoccupied catchments. A lack of continuity of macrophyte recording by the different RCS surveyors through the two catchments limits the possibility of using these data as an alternative tool for modelling alien species occurrence.

Chapter Four

Modelling using the River Corridor Survey Data

4.1 Aim and Introduction

The aim of the work presented in this chapter is to determine whether the crude RCS data on riparian habitats and riverine features could be used as a means of predicting either, the occurrence of an already well distributed alien species or, the potential range of such a species within an uncolonised catchment. Additionally such work, should it prove successful, would shed light upon the ecological requirements for the successful spread of these species.

After identifying the shortcomings of RCS data for modelling purposes and the solutions adopted, the work presented initially looks at the ability of such data to produce good predictive models. Resultant models are subsequently tested in their ability to predict aliens species occurrence on other areas of the same catchment and on different catchments. Models built using data from multiple catchments and using only sub-sets of variables are also created and their predictive ability assessed.

Riparian species modelling

Riparian plant species, despite being particularly easy to map, because of their linear network of occurrence and, in the case of hydrochorous species, their relatively simple dispersal mechanism, have received only scant attention with regards to modelling distribution and spread at a catchment scale. However, although the actual distribution of riparian species (e.g. Harper *et al.* 1997; Neiland *et al.* 1987; Pigott and Wilson 1978; Whitton 1998) and the spread of species at a larger spatial scale (e.g. Pysek 1991; Beerling *et al.* 1995) has been studied, there has been little work modelling the occurrence of such species in relation to local riparian variables. One of the few recent papers to model aquatic macrophytes (Narumalani *et al.* 1997) concentrates on large lakes in the USA and utilises physical features and not habitat based features. Similarly work in the USA (Toner and Keddy 1997) and Canada (Hill *et al.* 1998) uses physical parameters to predict riparian habitats. Models have been developed for predicting the occurrence of animals along riparian environments at a catchment scale however. These models mainly focus on bird (e.g. Buckton and Ormerod 1997; Rushton *et al.* 1994; Gregory *et al.* 1997) and fish species (e.g. Kruse *et al.* 1997) but also cover species as diverse as invertebrates (Richards *et al.* 1997) and mammals (Nadeau *et al.* 1995). One

of the most widely applied techniques for the production of such catchment models (most of which use large sets of variables), as well as non-riparian habitat-based models in the wider environment (e.g. Stowe *et al.* 1993; Green *et al.* 1994), is that of logistic regression.

Logistic regression modelling is now widely applied in ecological modelling, especially at larger spatial scales (e.g. Buckland & Ormorod 1997; Buckland *et al.* 1996) where databases compiled from vast arrays of variables can be easily assembled. Logistic modelling requires the dependent variable to be supplied as simple presence/absence data, which is appropriate for these large-scale studies. One problem, that was until recently frequently overlooked, is that of autocorrelation, which can occur when datasets are formed from contiguous, potentially inter-related areas. Several papers provide good summaries of the potential problems of autocorrelation (Legendre and Fortin 1989; Legendre 1993) and others provide adequate techniques to test for and to take account of these problems (Hinch *et al.* 1994; Smith 1994; Jarvinen 1992).

Results from work examining distribution patterns of plant species at large spatial scales emphasise the importance of factors such as climate, population density and geology (e.g. Pysek *et al.* 1998; Beerling *et al.* 1995; Huntley *et al.* 1995; Collingham *et al.*, in press). Such factors must necessarily also operate at the smaller scale of catchments but would be difficult to accurately assess at this small scale. Other variables chosen as good predictors at these smaller scales may however be operating as proxies for such large-scale variables at the catchment level; an obvious example being altitude acting as a proxy for climate. Such large-scale variables can then also be considered to be included in the catchment scale models. These models may therefore prove to be more informative than large-scale models, as they can account for the interactive effects of local habitat and physical characteristics with these large-scale distribution determinants.

4.2 Materials and Methods

To explore the potential for using Phase 1-type survey data (*sensu* JNCC 1993), such as is available from the RCS, as a means of predicting the potential occurrence of the three study species, logistic regression analyses were adopted as the means of analysis. The reduced variables listed in the previous chapter for the Tees and Wear catchments were used as independent variables, the RCS distribution data for each of the three species being used as the dependent variable. The RCS species distribution data were

augmented by the additional survey data from this research (see chapter 5). The variables were imported into the SPSS statistical package (Norusis 1994) to facilitate the production of logistic regression models. However before analyses could be undertaken manipulation of the data was necessary in order to avoid violating some of the basic assumptions of such statistical analyses. The major problem in using the RCS data was that of autocorrelation of the variables.

4.2.1 Autocorrelation: Testing and elimination

Spatial autocorrelation is a property that mapped data possess whenever they exhibit an organised pattern (Upton and Fingleton 1985) or, whenever there is systematic spatial variation in values across a map (Cliff and Ord 1981). Cliff and Ord (1981) define a lack of spatial autocorrelation as being when there is no connection between the variables (X_i, X_j) at any pair of regions (i,j) in the study area.

Associations in the RCS data

As an initial approach to examining the existence of autocorrelation of the three species between the left and right bank of sections of the River Wear simple 2x2 contingency tables were calculated and the distribution examined using a χ^2-test with Yates correction. These showed significant cross-bank association for the study species (Table 4.1)

Table 4.1: 2x2 tables of the occurrence of the study species on opposite banks of the River Wear.

a) *Impatiens*

Bank		Right	
		A	P
Left	A	90	6
	P	2	267

b) *Heracleum*

Bank		Right	
		A	P
Left	A	304	8
	P	3	50

c) *Fallopia*

Bank		Right	
		A	P
Left	A	265	40
	P	34	26

Separate 2x2 tables are shown for the presence (P) and absence (A) of a) *Impatiens* (χ^2 = 333.619, P<0.001), b) *Heracleum* (χ^2 = 283.216, P<0.001) & c) *Fallopia* (χ^2 = 29.482, P<0.001) along the left and right banks.

In addition to autocorrelations in the study species across the river, autocorrelative effects were also shown for the independent variables in adjacent and opposite sections. Because of such widespread occurrence of autocorrelation in the data it was not possible to use most standard statistical approaches on the unaltered dataset. Instead the data had to be manipulated to eliminate these effects.

Approaches to eliminate autocorrelation

a)The study species

One approach to eliminate autocorrelation is to calculate correlelograms to determine at what distance the effects are reduced to an acceptable level. The nature of the RCS data (2 rows by X columns) does not allow for the simple calculation of correlelograms so an alternative method of detecting a breakdown of autocorrelation was adopted. A χ^2 test was undertaken using the presence of a species in a section S and a section (S+x) upstream of section S for all reaches of the River Wear. At the point at which association between the sections breaks down it would be expected that a χ^2 test would indicate no association i.e. accept the null hypothesis.

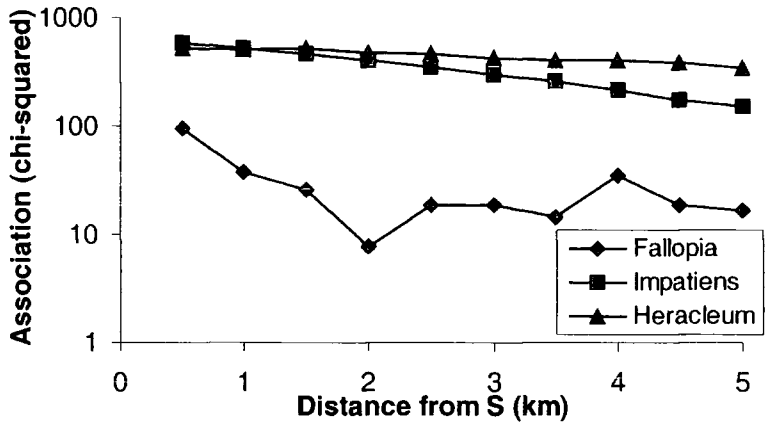

Figure 4.1 Results of 2x2 χ^2-tests comparing the occurrence of the three species in an RCS section, S, on the Wear catchment with its occurrence in a section (S + x) further upstream. n=730. Relevant χ^2 value for significance at P=0.01 is 6.635 (all values are significant).

Figure 4.1 indicates significant associations for all three species between their occurrence in a RCS section and their occurrence in all other sections on the same bank up to 5km upstream. This persistence of apparent autocorrelation is likely to be a function of the continuous occurrence of both *Heracleum* and *Impatiens* where they occur on the Wear catchment. Even for *Fallopia*, the species with the least continuous

distribution, the observed number of (1,1) occurrences is always greater than expected, even up to 5km (10 sections) away.

As it has been shown that autocorrelation cannot be easily eliminated from the distribution of the study species, an alternative means to counter these effects must be adopted. The method selected for this purpose is one suggested by Smith (1994) which is used in particular with reference to species distributions. Smith's method of accounting for this spatial information is to include it as an extra variable in the model, a proximity variable, rather than attempting to eliminate it, thereby taking account of the spatial patterning. For the analysis of the RCS data two proximity measures were calculated. One of these records the number of occurrences of the species in the adjacent cells upstream and downstream. The second is similar but includes the number of occurrence in the nearest three neighbours up- and down-stream of a cell. Subsequent analysis however found that the two were very strongly correlated. As the measure using only the two adjacent cells was always selected preferentially over the other, the latter was dropped from the analyses.

b) The RCS variables

The χ^2 analysis indicated that there is a strong autocorrelative effect between the two opposite banks of any reach of the catchments. It is therefore quite likely that the occurrence of a species and set of habitat criteria on one bank will lead to the occurrence of the species on the opposite bank. For this reason it was decided to combine the attributes of the left and right bank together to create a composite set of values for any reach of the rivers. This in turn would also serve to reduce autocorrelation in some of the bankside and river features. To further reduce any autocorrelative effects, especially in the bankside and land-use variables, only alternate reaches of a catchment were selected for subsequent analysis (the two resultant datasets being referred to as Tee1/Tee2 and Wea1/Wea2 hereafter for the Tees and Wear catchment respectively). This method has a secondary advantage in that a model created by this approach could then be tested on the remaining reaches of the catchment. This is analogous to Smith's (1994) approach of eliminating first order neighbours, hence reducing the effect of neighbouring cells reflecting similar environmental conditions (as mentioned by Augustin *et al.* 1996). The proximity measures were recalculated for these combined datasets.

4.2.2 Logistic regression methodology

With the data broken up into the four subsets (Tee1/Tee2 & Wea1/Wea2) and the effects of autocorrelation accounted for as much as possible, the building and testing of predictive logistic models could be undertaken.

As a standard approach the likelihood ratio (LR) method (Tabachnik and Fidell 1996) was adopted and, for most cases, the analyses were undertaken using firstly the full complement of available variables to build a particular model and then the forward LR and backward LR methods to remove redundant variables. As a standard rule the probability value for a variable to be entered in a regression model was set at 0.05, the probability value for removal being set at 0.1, with a maximum of 30 iterations and a constant always included in the model. From the three methods the most parsimonious model was selected on the basis of 1) the percent of present (P) and absent (A) records correctly predicted, 2) Cohen's κ statistic and 3) values of χ^2 compared to the degrees of freedom in the models. The subjective scale proposed by Prentice *et al.* (1992) was adopted as the means of assessing the κ statistic, with values assigned as in Table 4.2.

Table 4.2 Interpretation of Cohen's κ statistic as subjectively proposed by Prentice *et al.* (1992).

Cohen's κ	Model Fit
<0.4	Poor
0.4-0.55	Fair
0.55-0.7	Good
0.7-0.85	Very Good
>0.85	Excellent

The logistic regression models produce a predictive value for each cell (reach in the case of the catchments). These values are continuous between 0 and 1 with the cut-off point for predicting presence or absence arbitrarily set at 0.5. It may have been possible in some cases to improve the fit of models by altering this arbitrary value so as to maximise the fit of the model (e.g. Collingham *et al.*, in press; Huntley *et al.* 1995).

Because of the problem of not being able to use a proximity variable when testing predictive models elsewhere (such tests by definition should assume nothing of a species occurrence), models were built both with and without this variable to determine whether good models could be produced once it is excluded.

Altitude, being strongly correlated with many of the RCS variables and acting as a proxy for climate, is quite different to the other RCS variables. For this reason models were also built including and excluding this variable. Using these approaches it is possible to qualify the importance of autocorrelation and altitude in the predictive ability of these models.

During the process of running the analyses it was found on several occasions that the large number of variables that could be included in models led to the production of non-unique perfects fits for a species distribution. To eliminate such uninformative output, variables were excluded from the potential variable list and the model re-run. The variables chosen for exclusion were based on non-significant associations with the target species. This information was derived from the initial exploratory analyses of the variables (Chapter 3).

Initially individual models were built for each of the four data subsets and for each of the three species. The best such models for each dataset were then assessed by their ability to predict occurrences of the dependent species within the sections of the corresponding catchment not used in the model creation.

The best models for a particular catchment were then tested by their ability to predict occurrences in the whole of the alternative catchment. Similarly models were designed from a dataset created from a mixture of data from both catchments and subsequently tested on the remainder of the sections unused in the model creation.

To make the within-catchment predictions comparable, ideally the same reference data should be used to determine the occurrence of the study species i.e. the original RCS derived occurrences. These analyses were undertaken but are not displayed as the resultant models were almost identical to those created incorporating the updated Wear distribution data from the present survey. As might be expected the models derived from the initial Wear RCS distribution data were less good at predicting between catchments, though again similar variables were used in the different models.

One of the aims of this work was to determine whether occurrence/spread of the study species could be predicted using only simple variables. With this in mind the variables used in the building of these models were separated into four different categories in order to ascertain whether any of these variable subsets would provide equally good predictive models to those produced from the entire variable set. The four categories the variables were classified into were: bankside physical features, riverine features, bankside habitat/vegetation and surrounding land-use.

4.2.3 RCS macrophyte data

All of the modelling methods mentioned thus far have utilised only very basic biological habitat characteristics. The potential shortcomings of such generalised data are discussed later. To increase the biological resolution of such data it would be preferable to divide the habitats into categories of NVC (Rodwell 1991-99) or Braun-Blankuet style (Braun-Blaunkuet 1932) classifications. Such classification of habitat would be logistically prohibitive at a catchment scale and resultant models would be of limited application elsewhere because of a lack of other high-resolution data on which they could be tested. Attempts to classify the macrophytic vegetation recorded at the RCS 500m scale into NVC communities proved largely unsuccessful, probably largely because of the sections spanning several community types, both vertically and horizontally. Other limiting factors also included inadequate species recording in the RCS surveys and a paucity of coverage of such riverine vegetation types in the available NVC literature.

Logistic regression models were created using the RCS native macrophyte occurrences as independent variables with which to predict the study species distributions. The end models resulting from these analyses proved to be of only moderate predictive ability compared to those created using the habitat and physical parameters (see Table 4.3). An inherent problem with using these RCS macrophyte data was the variability in the species recording between different surveyors. This lack of standardisation rendered the RCS macrophyte data of negligible use in the creation of such models.

Table 4.3 Summary of logistic regression models using RCS macrophytes as predictor variables.

Species	Present	Absent	Overall	κ
Impatiens	75.64	80.31	78.08	0.56
Heracleum	37.84	98.06	88.9	0.46
Fallopia	10.32	98.84	83.56	0.14

If a modelling approach using other species as variables was to be adopted a more systematic set of distribution data would be needed.

4.3 Results

4.3.1 Modelling *Impatiens* on the Wear and Tees

Modelling within the Wear catchment

The forward LR regression produced the best model for predicting *Impatiens* using the Wea1 dataset (Table 4.4). Similarly for the Wea2 dataset the forward LR method was selected as producing the best-fit model (Table 4.5)

Table 4.4 Summary of the logistic regression results produced from the analysis of *Impatiens* on the Wear 1 dataset.

Variations	Model	χ^2	df	Classification (% correct)			κ
				P	A	Total	
Normal *	Full	110.792	17	95.05	67.44	88.59	0.66
Normal *	Forward	150.169	11	96.45	83.72	93.48	0.82
incl prox.	Forward	196.776	7	99.29	100	99.45	0.99
excl alt.	Forward	98.96	9	94.33	72.09	89.13	0.69

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001

Table 4.5 Summary of the logistic regression results produced from the analysis of *Impatiens* on the Wear 2 dataset.

Variations	Model	χ^2	df	Classification (% correct)			κ
				P	A	Total	
Normal *	Full	131.125	17	96.35	75	91.16	0.75
Normal	Forward	160.538	11	98.54	84.09	95.03	0.86
incl. prox	Forward	144.574	1	98.54	86.36	95.58	0.88
excl. alt	Forward	101.091	11	95.62	65.91	88.4	0.66

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001

Modelling within the Tees catchment

Table 4.6 Summary of the logistic regression results produced from the analysis of *Impatiens* on the Tees 1 dataset.

Variations	Model	χ^2	df	Classification (% correct)			κ
				P	A	Total	
Normal	Full	292.224	46	76.25	92.33	86.58	0.70
Normal	Forward	259.376	7	72.5	90.59	84.12	0.65
Incl prox.	Forward	320.348	5	81.25	91.99	88.14	0.74
Excl. alt	Forward	216.337	9	65.63	90.24	81.43	0.59

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001

Of the equations derived from the Tees datasets (Tables 4.6 & 4.7) the forward LR method provided the most parsimonious model for the Tees1 dataset and the backwards LR model for the Tees2 dataset

Table 4.7 Summary of the logistic regression results produced from the analysis of *Impatiens* on the Tees2 dataset.

Variation	Model	χ^2	df	Classification (% correct)			κ
				P	A	Total	
Normal	Full	342.845	46	84.46	92.75	89.86	0.78
Normal	Backward	324.617	17	82.43	92.39	88.92	0.75
incl. prox	Backward	374.606	15	86.49	94.57	91.75	0.82
excl. alt	Backward	276.696	17	75	90.94	85.38	0.67

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001.

The κ values shown in both Table 4.6 and Table 4.7 show a trend of increased goodness of fit when the autocorrelation variable is included and a decreased goodness of fit when altitude is excluded. These trends are mirrored in the *Impatiens* models on the Wear and emphasise the importance of altitude and the presence of *Impatiens* in proximate locations in determining whether the species is present at any point. In the Wear examples, the reduction in predictive power of the models arises due to a decreasing ability to correctly predict the absence of *Impatiens*. By contrast, the Tees examples show a decreasing ability to predict the species' presence.

The κ statistics for the most parsimonious models from the four datasets show that the normal models are very good predictive models and with altitude excluded they remain good or very good. When the proximity variable is included model fits improve still to be very good or excellent.

A summary of the significant variables used in the building of all of the *Impatiens* models is given in Table 4.8 overleaf. In one of these models none of the variables was deemed significantly different from zero according to the Wald statistic. However this statistic has an undesirable property in that, when the regression coefficient is large, the estimated standard error is too large (Norusis 1994). This is the case in this example where all coefficients are large and all variables are deemed insignificant. These variables are simply labelled as "x"s in the summary table as they cannot be assigned a level of significance.

Table 4.8 Significant variables used in the creation of the various *Impatiens* logistic regression equations.

Model	N	+pro	-alt	N	+pro	-alt	N	+pro	-alt	N	+pro	-alt
Dataset	We1	We1	We1	We2	We2	We2	Te1	Te1	Te1	Te2	Te2	Te2
Prox_var	n	x(+)	n	n	+++	n	n	+++	n	n	+++	n
Constant	+++	x(-)		++	---	-	---	---	---	--	--	---
Gr/he	++		+			+++					--	-
Altitude	---		n	---		n	---	--	n	---	-	n
Arable											++	
Am_mo_ga				-								
Bed			-			---						
Bould	+++								--	++		
Cobble							+++	+++	+++			
Decid ba				+		++	++	++	+++			++
Decid lu		x(+)	+++			+++				++		++
Derelict				-						-		
Disturb		x(-)	-			---						
Farming	--										-	--
He/ru			++	-		++	+++	+++	+++	+++	++	+++
Human	+									-		
Islands	+		++									
Min_slop										-	-	-
Max_slop										+++	++	+
Max_wid	---				--		+++		+++	++	++	+++
Min_dep									+	+++	++	+++
Min_ht										++	++	++
mixed		x(+)				+						
Mud	-											+
Road		x(-)				+						
Road_br	---											
Shingle		x(+)				-	+					
Sh_mo										+++	++	+++
Silt			++			++	---		---	---	---	---
Sof_slum	-		-	-					+++			+
Stream												
Tot wood	+++											
Urban		x(+)								+		++
Unman				++		++						

The normal model, the model including the proximity variable and the model excluding altitude are labelled N, +pro and –alt respectively. Significance levels of 0.05, 0.01 and 0.001 are respectively represented as +, ++ and +++; or -, -- and --- depending on coefficient sign. For variables labelled "x" see text.

The summary table for the most parsimonious models reassuringly highlights a trend of the most important variables being components of most models. The models including a proximity variable are seen to require few variables other than the proximity variable for good predictive abilities. Models without the proximity variable utilise altitude as a very important variable along with river width and, to a lesser extent, herb/ruderal vegetation, silt, soft/slumping banks and woodland. With altitude also excluded herb/ruderal vegetation and silt become very important explanatory variables as well as river width and the deciduous woodland categories. Other less consistently important variables include cobbles, minimum river depth and soft/slumping banks.

Testing *Impatiens* models within catchments

Testing the most parsimonious models created from one half of the catchment data on the remaining reaches not used in the model creation (Table 4.9) generally resulted in good predictive abilities (from κ statistics). With altitude removed the models predictive abilities dropped but still bordered on good to fair.

Table 4.9 Test of the ability of the most parsimonious model for a dataset to predict the occurrence of *Impatiens* in the remaining sections of a catchment.

Model data	Test data	Include altitude?	Classification (% correct)			κ
			P	A	Total	
Tee1	Tee2	Y	72.3	88.77	83.02	0.62
Tee1	Tee2	N	68.92	86.96	80.66	0.57
Tee2	Tee1	Y	65.63	90.59	81.66	0.59
Tee2	Tee1	N	63.75	87.11	78.75	0.52
Wea1	Wea2	Y	86.86	77.27	84.53	0.6
Wea1	Wea2	N	85.4	63.64	80.11	0.48
Wea2	Wea1	Y	92.91	81.4	90.22	0.73
Wea2	Wea1	N	90.07	67.44	84.78	0.58

Results are shown for models both including and excluding the altitude variable.P<0.0001 for all models.

Testing *Impatiens* models between catchments

The results of using the best models for each of the two catchments to predict distributions on the whole of their corresponding alternate catchment are summarised in Table 4.10. These show that, with the exception of the model built from the Tee1 data and including altitude, none of the most parsimonious models can predict occurrence of *Impatiens* in another catchment; κ values for these predictions being very poor for all but the noted exception. It therefore seems that models built on only one catchment using such crude data cannot accurately predict distribution patterns elsewhere. Figure 4.2 shows an example of cross-catchment predictions and highlights the areas where shortfalls in predictive ability are occurring.

Table 4.10 Test of the ability of the most parsimonious model for a catchment to predict the occurrence of *Impatiens* in the whole of the other catchment.

Model data	Species	Include altitude?	χ^2	df	Classification (% correct)			κ
					P	A	Total	
Tee1	*Impatiens*	Y	259.4	7	72.5	90.59	84.12	0.65
Tee1	*Impatiens*	N	216.3	9	30.94	71.26	40.55	0.13
Wea2	*Impatiens*	Y	160.5	11	20.07	89.29	44.55	0.07
Wea2	*Impatiens*	N	101.1	11	13.5	91.23	40.99	0.04

Results are shown for models both including and excluding the altitude variable. P<0.0001 for all models.

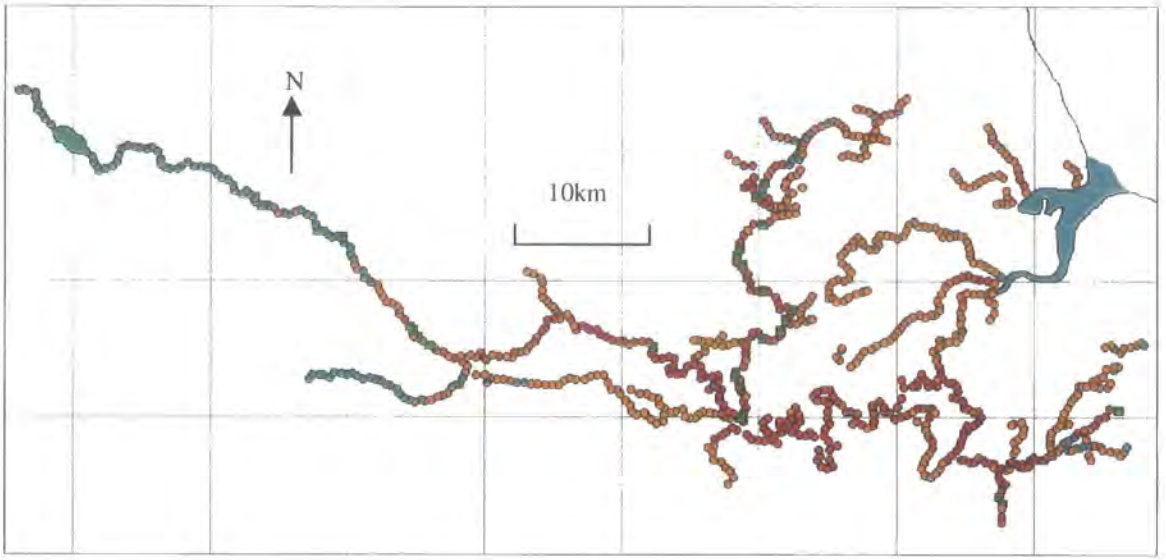

Figure 4.2 A comparison of the observed and predicted distributions of *Impatiens* along the Tees catchment, using the best Wear model to predict occurrence. Correctly predicted presence is shaded red, and absence cyan. Observed sites of absence which the model predicts as occupied are shaded orange, and observed presence but predicted absence are shaded green.

Amalgamated models

To determine whether models could be produced which were accurate predictors and not peculiar to one catchment, a dataset was constructed using part of both the Tees and Wear catchments and including altitude. The summaries of the resultant models and their ability to predict occurrences on unused reaches are shown in Table 4.11.

From the summary variables list (Table 4.12) it can be seen that the most important variables in these composite models are also those common to the previous single catchment models, with the exception that woodland is largely excluded. The common variables also consistently have the same coefficient sign in the various models. The κ statistics in Table 4.11 all indicate good model fits in predicting occurrences in test reaches, though the inability to test these models on an unrelated catchment limits our ability to predict their wider application. An example of predictions using an amalgamated catchment model is shown in Figure 4.3 and highlights the good model fit.

Table 4.11 Summary of the ability of the models build using one half of the Wear and one half of the Tees catchments to predict the occurrence of *Impatiens* in the remaining sections of the two catchments.

Model data	LR method	Model/ Test data	χ^2	df	Classification (% correct)			κ
					P	A	Total	
TeWe1	Full	M	406.1	45	84.72	85.15	84.94	0.70
TeWe1	Full	T	-	-	82.81	83.44	83.14	0.66
TeWe1	Forward	M	345.5	11	79.73	83.94	81.93	0.64
TeWe1	Forward	T	-	-	81.4	81.88	81.65	0.69
TeWe1	Backward	M	392.3	24	81.73	84.85	83.36	0.66
TeWe1	Backward	T	-	-	82.46	83.44	82.98	0.66

Results indicate the original model created (M) and beneath, the ability of that model to predict occurrence in the other sections (T). P<0.0001 for all models.

Table 4.12 Summary of the significant variables in the forward and backward models built to predict the occurrence of *Impatiens* using the composite dataset.

Variable	Forward	Backward	Variable	Forward	Backward
Gr/he	+++	+++	Road_br		--
Altitude	---	---	Silt	---	---
Arable	+		Sh_mo		-
Artific		+	Soft_slu	+	
Conif_ba	+	+++	He/ru	+++	+++
Dep_min	+	+	Tot_wood	+	
Unman	+	++	Wild	-	-
Islands		++	Constant	---	-
Mud		-			

Significance levels of 0.05, 0.01 and 0.001 are respectively represented by +, ++ and +++; or -, -- and --- depending on coefficient sign.

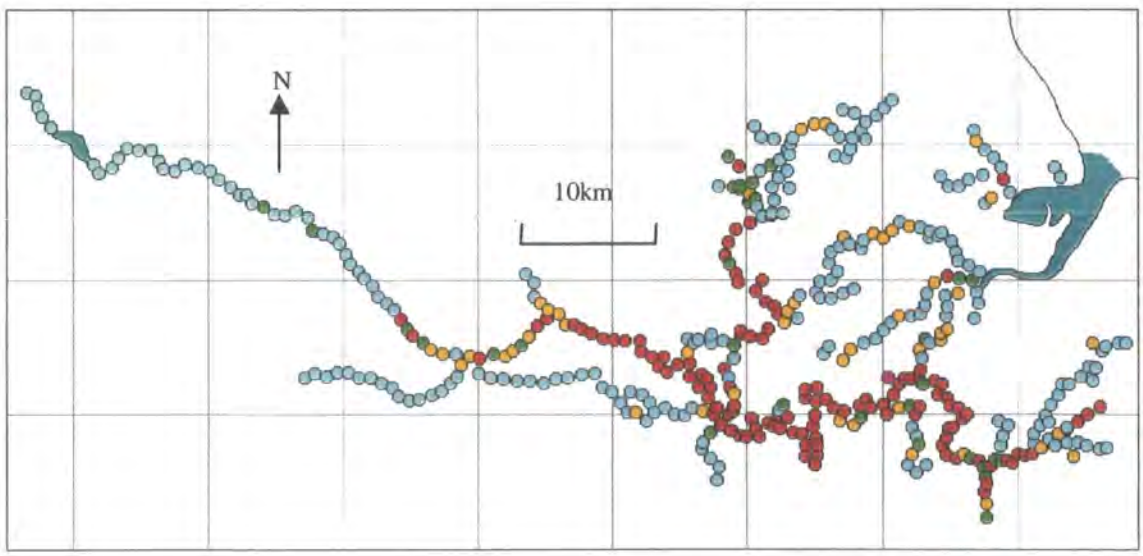

Figure 4.3 A comparison of the observed and predicted distributions of *Impatiens* along the Tees catchment, using the amalgamated Tees/Wear model to predict occurrence. Colours are as in Figure 4.2.

4.3.2 Modelling *Heracleum* on the Wear and Tees

Modelling within the Wear catchment

The regression equations produced using the Wea1 dataset are summarised in Table 4.13. The backwards LR method provided the most parsimonious model for both this dataset and also the Wea2 dataset (Table 4.14). However in the latter datset with the proximity variable included non-unique perfect fits are produced, even with a much reduced variable list.

Table 4.13 Summary of the logistic regression results produced from the analysis of *Heracleum* on the Wear 1 dataset.

Variations	Model	Chi-2	df	Classification (% correct)			κ
				P	A	Total	
Normal *	Full	114.711	10	81.25	97.37	94.57	0.81
Normal *	Backward	111.788	5	81.25	98.03	95.11	0.82
Incl prox *	Backward	143.763	4	93.75	99.34	98.37	0.94
excl. alt *	Backward	50.594	4	34.38	95.39	84.78	0.36

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001.

Table 4.14 Summary of the logistic regression results produced from the analysis of *Heracleum* on the Wear 2 dataset.

Variations	Model	Chi-2	df	Classification (% correct)			κ
				P	A	Total	
Normal *	Full	139.838	14	90	98.01	96.69	0.88
Normal *	Backward	139.22	12	90	97.35	96.13	0.86
Incl prox *	Backward	-	-	100	100	perfect	-
excl. alt *	Backward	77.138	9	56.67	95.36	88.95	0.57

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001.

The summary statistics of the regression models using the Wear1 and Wear 2 datasets (Tables 4.13 & 4.14) highlight the importance of adjacent areas containing *Heracleum* and the effects of altitude in predicting presence of the species. In fact, when the proximity variable is included very few, if any, other variables are needed to produce a very accurate model. This is due to the single continuous distribution of the species on the Wear. The normal models are also very good predictors (based on κ values) of occurrence and with altitude excluded, resultant models are still good. The inclusion of the proximity variable results in an increase in overall predictive ability whilst the exclusion of altitude causes a marked reduction in the ability of the models to predict

the occurrence of *Heracleum* (the ability of the models to predict absence remaining high). The changes in the κ statistic due to removing altitude are most marked in the Wear1 dataset.

Modelling within the Tees catchment

Of the equations derived from the Tees1 dataset the backward LR method proved most parsimonious using the normal variable set (Table 4.15). Similarly the backwards LR method also produced the most parsimonious model for the Tees2 datset (Table 4.16).

Table 4.15 Summary of the logistic regression results produced from the analysis of *Heracleum* on the Tees1 dataset.

Variations	Model	Chi-2	df	Classification (% correct)			κ
				P	A	Total	
Normal *	Full	205.545	29	87.18	99.02	97.99	0.87
Normal *	Backward	198.111	11	82.05	98.77	97.32	0.83
incl prox. *	Backward	198.111	11	82.05	98.77	97.32	0.83
excl. alt *	Backward	161.895	11	66.67	98.04	95.3	0.69

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001.

Table 4.16 Summary of the logistic regression results produced from the analysis of *Heracleum* on the Tees2 dataset.

Variables Excluded	Model	Chi-2	df	Classification (% correct)			κ
				P	A	Total	
Normal *	Full	198.991	29	77.27	97.89	95.75	0.77
Normal*	Backward	184.039	6	77.27	96.84	94.81	0.73
Incl prox.*	Backward	184.04	6	77.27	96.84	94.81	0.73
excl. alt	Backward	204.123	17	84.09	98.42	96.93	0.83

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001.

The previously observed trends from the Wear data of increasing κ statistic values when the proximity variable is included and decreasing κ statistic values when altitude is excluded are not apparent on the Tees. This may indicate that on the Tees altitude and nearby seed sources are less important than on the Wear. This is likely to be a function of the less continuous distribution of *Heracleum* in the Tees lowlands. This in itself could simply be a function of the Tees distribution data relying solely on RCS information; additional surveying on the Wear filled in a lot of gaps in the distribution of *Heracleum*. Tetrad distribution maps from Graham (1988) also indicate *Heracleum* to be more widely distributed on the Tees than the RCS suggests.

A summary of the significant variables in the various *Heracleum* models is given in Table 4.17. Models from the Tees datasets show that, unlike the Wear models, the proximity variable is not the most important variable in predicting the species occurrence along the Tees. Instead altitude, river width and disturbance are more important.

Table 4.17 Significant variables used in the creation of the various *Heracleum* logistic regression equations.

Model	N	+pro	-alt	N	+pro	-alt	N	+pro	-alt	N	+pro	-alt
Dataset	We1	We1	We1	We2	We2	We2	Te1	Te1	Te1	Te2	Te2	Te2
Prox_var	n	+++	n	n	+++	n	n		n	n	+	n
Constant		---	---	+	---	---		-	---		---	
Altitude	---		n	--		n	---	---	n	---	---	n
Arable												+
Bed												-
Bould			--			--						
Cobbles						++						
Decid_lu						+			+++			
Disturb						---	--	--	---	-	-	---
Farming				+								
Grass												-
Gr/he							--	--		-	-	--
He/ru									++			
Max_wid	-		+++			+++	+++	+++	+++	+++	+++	+++
Min_slop				-					---			
Max_slop									---			---
Mud												+
Reed				-								
Tot_res												---
Road		+				+++						
Sh_mo			+			++						
Silt				+		-			++			
Soft_slu	-			-		--						
Urban												++
Tot_wood	+		+	+								
Shingle									++			

The normal model, the model including the proximity variable and the model excluding altitude are labelled N, +pro and –alt respectively in the model title row. Significance levels of 0.05, 0.01 and 0.001 are respectively represented as +,++ and +++ or -, -- and --- depending upon the coefficient sign.

Altitude, river width, woodland and the disturbance variables are included in almost all equations as important factors in determining distribution; woodland and river width with positive coefficients and altitude and disturbance with negative.

In models without the proximity variable, altitude becomes the most important explanatory variable with river width and woodland of lesser importance. Other variables such as disturbance, grass/herb vegetation and soft/slumping banks also appear in these models.

With altitude excluded from the variable list, river width and woodland become increasingly important; river width perhaps acting as a proxy for altitude. Disturbed/derelict areas are also included in the models, acting to reduce the likelihood of *Heracleum*'s occurrence in an area. Other variables included are boulders, bank slope and smaller substrates (silt, mud & shingle).

Most of the significant variables in the different models tend to have fairly similar weighting in any model compared to the other variables. The only regular exception to this rule is the constant, which often has a coefficient value higher than that of the other variables.

Testing *Heracleum* models within catchments

With altitude included, all of the models are capable of producing good fits (from the κ interpretation) when tested on unused reaches in the same catchment (Table 4.18). When altitude is excluded from the variable lists the models are only fair in their predictions. This emphasises the importance of altitude in predicting *Heracleum* in both catchments.

Table 4.18 Test of the ability of the most parsimonious model for a dataset to predict the occurrence of *Heracleum* in the remaining sections of a catchment.

Model data	Test data	Include altitude?	Classification (% correct)			κ
			P	A	Total	
Tee1	Tee2	Y	70.45	97.37	94.58	0.70
Tee1	Tee2	N	52.27	96.32	91.75	0.52
Tee2	Tee1	Y	69.23	96.81	94.41	0.65
Tee2	Tee1	N	58.97	95.34	92.17	0.53
Wea1	Wea2	Y	83.33	97.35	95.03	0.82
Wea1	Wea2	N	66.67	92.05	87.85	0.57
Wea2	Wea1	Y	71.88	94.08	90.22	0.66
Wea2	Wea1	N	53.13	91.45	84.78	0.46

Results are shown for models both including and excluding the altitude variable. $P<0.0001$ for all models.

Testing *Heracleum* models between catchments

From the most parsimonious models the best two were selected from each catchment, one including altitude and one excluding it. Each of these four models were then tested by their ability to predict the occurrence of *Heracleum* on the whole of the alternate catchment. The results of these tests are displayed in Table 4.19. These results display two main findings: firstly that none of the four models prove other than a poor predictor for *Heracleum* in the alternate catchment and secondly, that the Wear models seem to be better predictors than the Tees models.

Table 4.19 Test of the ability of the most parsimonious model for a catchment to predict the occurrence of *Heracleum* in the whole of the other catchment.

Model data	Species	Include altitude?	Chi-2	df	Classification (% correct)			κ
					P	A	Total	
Tee1	*Heracleum*	Y	198.1	11	9.68	95.71	81.1	0.07
Tee2	*Heracleum*	N	203	17	4.84	96.37	81.82	0.02
Wea1	*Heracleum*	Y	111.8	15	60.24	80.96	78.99	0.25
Wea2	*Heracleum*	N	92.4	10	33.73	86.93	81.86	0.16

Results are shown for models both including and excluding the altitude variable. $P<0.0001$ for all models.

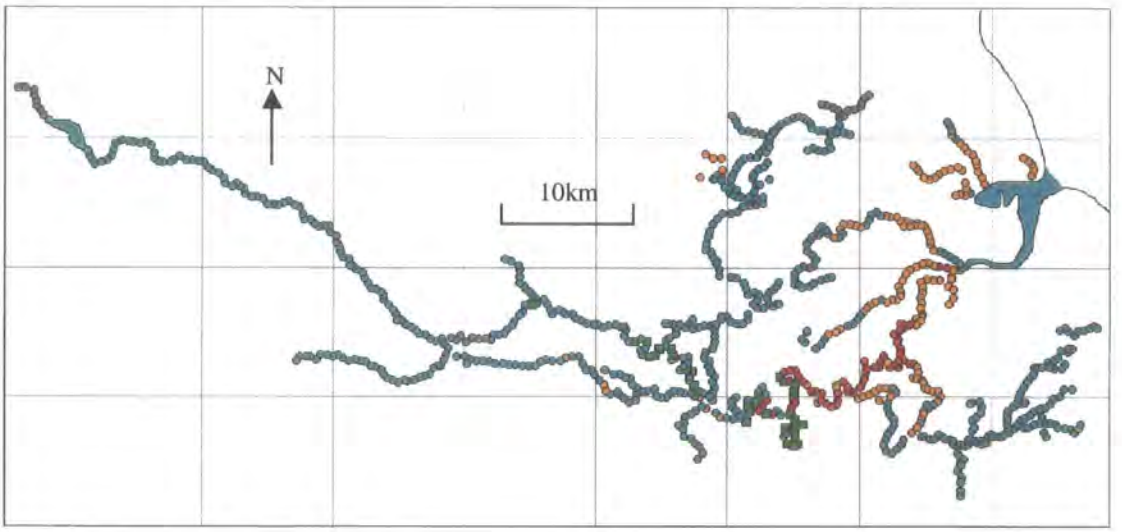

Figure 4.4 A comparison of the observed and predicted distributions of *Heracleum* along the Tees catchment, using the best Wear model to predict occurrence. Correctly predicted presence is shaded red, and absence cyan. Orange circles show observed sites of absence which the model predicts as occupied and green squares show observed presence but predicted absence.

Figure 4.4 gives an example of a predicted distribution of *Heracleum* on the Tees, using a model derived from the Wear catchment. This highlights lowland tributaries being incorrectly predicted to support *Heracleum* and actual upstream occurrences being unpredicted.

Amalgamated *Heracleum* model

As a result of the poor cross-catchment predictive abilities shown in the previous section, amalgamated catchment models were created for *Heracleum* and tested on unused reaches in the same way as was done for *Impatiens* (see 4.3.1). The results of these tests are summarised in Table 4.20. These results indicate no difference in the ability of the three test methods, all yielding similar κ values. As with *Impatiens*, these composite models prove much better at predicting occurrences across both catchments than those created from single catchment databases. All the models have good

predictive capabilities, κ statistics being only slightly lower than for the within-catchment predictions (Table 4.18). Figure 4.5 compares the predictions of an amalgamated catchment model with the actual distribution of *Heracleum* on the Tees, highlighting the good predictive fits that can be achieved.

Table 4.20 Summary of the ability of the models build using one half of the Wear and one half of the Tees catchments to predict the occurrence of *Heracleum* in the remaining sections of the two catchments.

Model data	Test method	Model/ Test data	Chi-2	df	Classification (% correct)			κ
					P	A	Total	
TeWe1	Full	M	290.82	45	71.83	97.32	94.45	0.71
TeWe1	Full	T	-	-	62.16	95.86	91.74	0.60
TeWe1	Forward	M	262.22	7	70.42	97.32	94.29	0.70
TeWe1	Forward	T	-	-	62.16	96.42	92.23	0.62
TeWe1	Backward	M	268.45	9	71.83	98.04	95.09	0.74
TeWe1	Backward	T	-	-	63.51	96.61	92.56	0.63

Results indicate the original model created (M) and beneath, the ability of that model to predict occurrence in the other sections (T). P<0.0001 for all models.

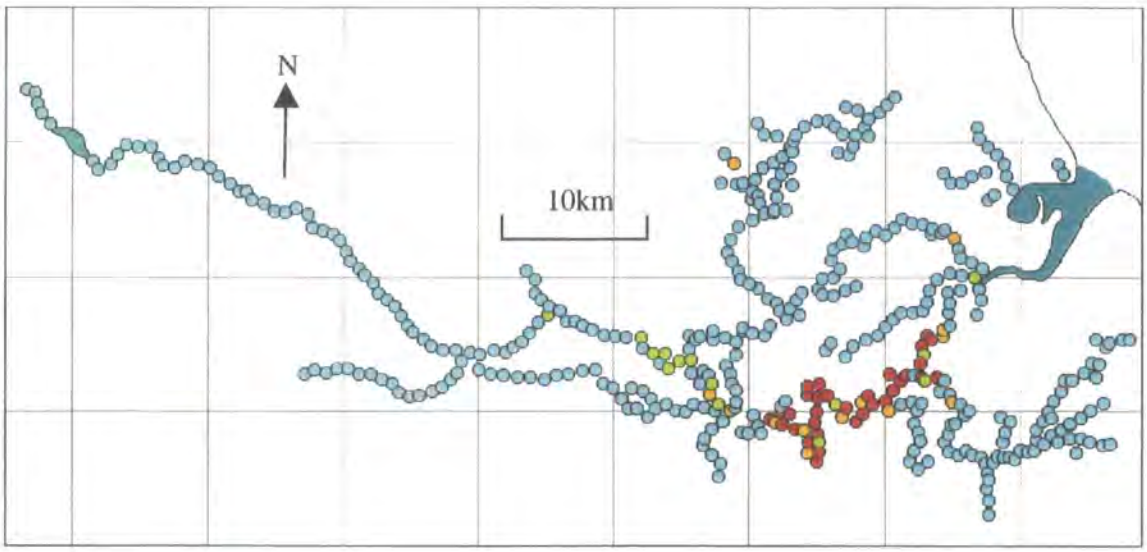

Figure 4.5 A comparison of the observed and predicted distributions of *Heracleum* along the Tees catchment, using the amalgamated Tees/Wear model to predict occurrence. Red and cyan circles represent correctly predicted presence and absence respectively. Orange circles represent sites of observed absence but predicted presence and green circles represent sites of observed presence but predicted absence.

As with the *Impatiens* models, when the mixed catchment models and the various single catchment models are compared similarities in variables are apparent (Tables 4.17 and 4.21). Among the important variables common to both sets of models are altitude, woodlands, river width and disturbance. The common variables consistently have the same coefficient sign, with the disturbance variable, altitude and the constant having

negative coefficients. The only variables that were included in several of the within-catchment models but which are excluded here are the substrate variables

Table 4.21 Summary of the significant variables in the forward and backward models built to predict the occurrence of *Heracleum* using the composite dataset.

Variable	Forward	Backward
Constant	---	---
Altitude	---	---
Cobble	++	+
Cobif_ba	+	+
Disturb	---	-
Reed	+	
Tot_wood	+++	+++
Wid_max	+++	+++

Significance levels of 0.05, 0.01 and 0.001 are respectively represented by +, ++ and +++ or -, -- and --- depending on the coefficient sign.

4.3.3 Modelling *Fallopia* on the Wear and Tees catchments

Because of a paucity of occurrences of *Fallopia* on the Tees (there were only six sites recorded) a decision was made not to use the Tees data for the creation or testing of models but instead to use the Wear catchment data only.

Modelling on the Wear catchment

The most parsimonious models for predicting *Fallopia* distribution using the Wear1 dataset are summarised in Table 4.22. The backward LR method provided the most parsimonious models for both of the Wear datasets. When altitude was omitted from the variable lists for the two datasets the resultant models were in fact identical to those including it.

Table 4.22 Summary of the Logistic regression results produced from the analysis of *Fallopia* on the Wear 1 dataset.

Variations	Model	Chi-2	df	Classification (% correct)			κ
				P	A	Total	
Normal	Full	61.244	46	56.14	88.98	78.80	0.48
Normal	Backward	61.244	6	36.84	90.55	73.91	0.31
Incl prox.	Backward	52.015	3	42.11	91.34	76.09	0.37
excl. alt	Backward	33.819	6	36.84	90.55	73.91	0.31

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001.

The most parsimonious equations using the Wear2 dataset are shown in Table 4.23 (overleaf).

Table 4.23 Summary of the Logistic regression results produced from the analysis of *Fallopia* on the Wear 2 dataset.

Variations	Model	Chi-2	df	Classification (% correct)			κ
				P	A	Total	
Normal	Full	66.354	46	53.06	92.42	81.77	0.60
Normal	Backward	40.5	8	40.82	91.67	77.90	0.37
Incl prox.	Backward	53.497	6	51.02	89.39	79.01	0.43
excl. alt	Backward	40.5	8	40.82	91.67	77.90	0.37

"Variations" indicates analyses where the proximity variable was included (incl. prox) or altitude excluded (excl. alt.) Asterisks in this column indicate that a reduced variable list was used to create the model (see 4.2.2 Methods). All models are significant at the level of 0.0001.

The classification data and the κ statistics for the *Fallopia* models using these two datasets both indicate much poorer predictive powers than those shown previously for *Heracleum* and *Impatiens*. The *Fallopia* models prove particularly weak in their ability to predict *Fallopia*'s presence, remaining good at predicting absences. There is a slight increase in the κ statistic when the proximity variable is included, indicating that the occurrence of *Fallopia* in adjoining areas is of some importance in the distribution of *Fallopia* but not to the same extent as for *Heracleum* and *Impatiens*.

The best fit model for *Fallopia* with the proximity variable included borders on being a fair model, according to the κ statistic, but the two models excluding it are both poor predictors, even when tested on data used in the models creation. The proximity variable is consistently the most important of the variables in the models for the Wear when included (Table 4.24).

When the proximity variable is excluded the only variable common to both models is that of soft/slumping banks, with a negative co-efficient. Several classes of small substrate were selected as variables with positive coefficients (silt, mud & sand) but none are common to both models and only mud is individually significant. Other significant variables from the models include woods and residential areas (with positive coefficients) and rock/hard earth cliffs (with a negative coefficient). Generally though there are few variables common to the models created.

Table 4.24 Significant variables used in the creation of the various *Fallopia* logistic regression equations.

Model	Nor	+pro	+Nor	+pro
Dataset	We1	We1	We2	We2
Prox_var	n	+++	n	+++
Constant	---	---		-
Bridge		++	--	
Decid_lu	+			
Hard_cls	-			
Max_slop			-	
Min_ht			-	
Mud			++	++
Tot_res	++			
Road				-
Soft_slul	--	--	--	

The normal model, the model including the proximity variable are labelled N and +pro respectively in the model title row. Significance levels of 0.05, 0.01 and 0.001 are represented as +, ++ and +++ or -, --, --- respectively; depending on coefficient sign.

Testing models within and between catchments

Because of the few sites of occurrence of *Fallopia* on the Tees catchment it was decided not to use the Wear models to predict distributions on the Tees. Any results that did arise could not have been easily explained with such a small number of sites. Additionally, as models were not created for *Fallopia* on the Tees, no tests of models between catchments were undertaken. This section therefore deals solely with within catchment predictions from the Wear data.

Table 4.25 Test of the ability of the most parsimonious model for a dataset to predict the occurrence of *Fallopia* in the remaining sections of a catchment.

Model Data	Test data	Include altitude?	Classification (% correct)			κ
			P	A	Total	
Wea1	Wea2	Y	26.53	83.33	67.96	0.11
Wea2	Wea1	Y	15.79	82.86	61.96	0.01

Results are shown for models both including and excluding the altitude variable.

As might be expected from the poor model fits to the data use in their creation, the ability of these models to predict occurrence on the remaining sections of the catchment was very limited (Table 4.25). Even so, the extent of this inability to predict occurrence is greater than might be expected, suggesting that these poor models may not be using sound ecological variables which could be applied elsewhere. Similar results could be expected had tests between catchments been undertaken.

Amalgamated *Fallopia* models

Amalgamated models using data from both catchments (see Table 4.26 for summary) are no better in their predictive powers than those built from single catchment data, though several of the significant variables are common to both these models (Table 4.27) and may therefore be based on general ecological requirements.

Table 4.26 Summary of the ability of the models build using one half of the Wear and one half of the Tees catchments to predict the occurrence of *Fallopia* in the remaining sections of the two catchments.

Model Data	Test method	Model/ Test data	Chi-2	df	Classification (%correct)			κ
					P	A	Total	
TeWe1	Full	M	145.02	45	32.26	98.07	91.60	0.39
TeWe1	Full	T	-	-	19.30	95.44	88.26	0.18
TeWe1	For	M	109.2	10	17.74	98.24	90.33	0.23
TeWe1	For	T	-	-	21.05	96.53	89.42	0.22
TeWe1	Back	M	130.46	16	22.58	98.24	90.81	0.29
TeWe1	Back	T	-	-	19.30	95.44	88.26	0.18

Results indicate the original model created (M) and beneath, the ability of that model to predict occurrence in the other sections (T). P<0.0001 for all models.

Table 4.27 Summary of the significant variables in the forward and backward models built to predict the occurrence of *Fallopia* using the composite dataset.

Variable	Forward	Backward
Constant	---	---
Gr/he	++	++
Cobbles	+++	+++
Conif_ba	++	+
Decid_lu	+	++
Grass	--	-
Hard_cls	--	--
Islands		+
Mixed	-	--
Reed		+
Roads		+
Sand	++	++
Silt	--	-
Tot_res		++

Significance levels of 0.05, 0.01 and 0.001 are represented by 1,2 and 3 respectively. Variables with a negative co-efficient are in bold type.

If the various models for *Fallopia* are compared it can be seen that there are a few variables common to most models (Table 4.24 & 4.27). Those variables common to both the within-catchment and between-catchment models are deciduous woodland, solid earth/rock cliffs, roads and residential/urban areas.

4.3.4 Modelling species distributions using subsets of the RCS database

The results of modelling species distribution using only subsets of the entire range of variables available are summarised in Table 4.28. From the κ values it is apparent that none of the variable subsets is capable of producing a good model for any of the study species. The results do show however that different subsets of variables are better predictors for different species. This may be indicative of the fact that different features are more important for the different species.

Table 4.28 Summaries of the predictive abilities of models created using only subsets of the available variables.

Variable list	Species	Chi-2	df	p	Classification (% correct)			κ
					P	A	Total	
Land-use	*Impatiens*	29.08	3	0.0001	100	0	76.63	0
Riverine	*Impatiens*	32.75	4	0.0001	95.04	25.58	78.80	0.26
Bankside	*Impatiens*	45.5	6	0.0001	93.62	39.53	80.98	0.38
Habitats	*Impatiens*	41.98	3	0.0001	92.91	46.51	82.07	0.44
Land-use	*Heracleum*	31.77	4	0.0001	31.25	96.05	84.78	0.34
Riverine	*Heracleum*	34.21	2	0.0001	21.88	94.74	82.07	0.21
Bankside	*Heracleum*	19.09	2	0.0001	9.38	98.68	83.15	0.12
Habitats	*Heracleum*	10.51	2	0.0052	3.13	100	83.15	0.05
Land-use	*Fallopia*	10.93	2	0.0042	14.04	95.28	70.11	0.12
Riverine	*Fallopia*	-	0	0.0001	0	100	69.02	0
Bankside	*Fallopia*	16.01	1	0.0001	0	100	69.02	0
Habitats	*Fallopia*	-	0	0.0001	0	100	69.02	0

The different subset categories used are surrounding land-use, riverine variables, bankside physical parameters and bankside habitats. The data used to create the models was that from the Wear catchment and the method adopted was forward LR for all runs.

With *Impatiens* it seems that the bankside habitats and physical features are the most important predictor variables. This may be expected for a well-distributed species that is capable of reaching most areas of the river and is therefore limited most by habitat suitability. Conversely *Heracleum* is best predicted by surrounding land-use and riverine variables. Surrounding land-use may reflect potential sites of original introductions whereas riverine variables may reflect suitable seed deposition environments. Similarly surrounding land-use is the only subset that provides any predictive ability for *Fallopia* and may again reflect sites of original introductions.

The variables included in these subset models generally reflected the important variables selected for the full models for each species as summarised in earlier sections.

4.4 Discussion

Best-fit models for *Impatiens* and *Heracleum* produced from single catchment data vary from excellent to very good depending upon the variables included. The reductions in predictive ability of the models as the proximity variable and the altitude variable are excluded highlight their importance in determining the distribution of these two species. The proximity variable, altitude variable and river width variables alone can account for about 90%, 80% and 70% respectively of the occurrence of the two species. Other important variables for both species include woodland and substrate types, and additionally herb/ruderal vegetation for *Impatiens*, and disturbed areas for *Heracleum.* Important variables are consistently chosen in the models for the species and most can be rationally explained based on individual species ecology.

The fact that the various models produced for each species all have their most important variables in common and with the same coefficient signs is reassuring, suggesting that such modelling techniques can extract meaningful variables. The inability of a model produced on one catchment to predict a species occurrence on another catchment is more puzzling. One possible explanation for this is that species distributions on the catchment from which the model was created are not yet in equilibrium with the habitat availability, thereby overlooking some important factors in the species distribution. Conversely, the species may not yet be well dispersed on the test catchment and hence the model will over-represent the species occurrence. In the latter case it would be expected that the model would predict more occurrences than actually occur but that the predictions of absence would still remain quite good (perhaps Figure 4.4 shows an example of this). The evidence from limited species range expansion, earlier catchment surveys (Graham 1988, Whitton unpublished) and the long duration of *Impatiens* occurrence in the two catchments (Graham 1988) would however suggest that this species at least is in equilibrium with its surroundings and, in most cases, is not limited by dispersal. *Heracleum* however, being a later colonist of the Tees and Wear than *Impatiens* and with a longer generation time, may still be spreading within either or both of the two catchments. It is difficult to say whether the large increase in sites on the Wear, arising from surveys carried out during the present research (Chapter 5), is an artefact of poor initial recording or real range expansion but it is possibly a combination of the two.

When the patterns of observed and predicted occurrence between catchments were examined for *Heracleum* and *Impatiens* (Figures 4.2-4.5), major discrepancies were apparent. For *Impatiens* the Wear model over-predicts the sites of occurrence on the Tees. Most of these sites are small coastal tributaries but they also include the upper regions of the River Skerne and the lower sections of Clow Beck. The fact that all these erroneous sites are narrow reaches suggests that the use of river width as a driving variable with a negative coefficient on the Wear may be to blame. It could also be due to inadequate sampling of coastal tributaries in the original Wear RCS but equally may result from dispersal limitations or even under-recording.

For the *Heracleum* model altitude is the major driving variable. This seems to cause incorrect selection of lowland tributaries as sites of occurrence on the Tees and the omission of upland sites. As for *Impatiens,* coastal tributaries may be uncolonised due to dispersal limitations.

It may also be that the models created from only one catchment put too much emphasis on using variables which are typical of a zone in which a species occurs on that catchment but which will not necessarily be the same for other catchments. Variables that could be included in this category include river width, river depth and perhaps substrate types.

For *Heracleum* an alternative explanation for poor cross-catchment modelling may be that the species has been limited in its upstream occurrence by a lack of anthropogenic introductions higher up the catchments of one or both rivers. *Heracleum* may also be limited by factors not covered by the available RCS variables, for example a need for soils of a particular depth, moisture content or nutrient status. With its dominant growth pattern and fairly catholic riparian habitat selection it may be difficult to model using the available habitat variables. Having said this, its ability to be predicted both within and between catchments is very similar to that of *Impatiens* so perhaps small sample range (in terms of number of catchments) in building the models is the problem for both species.

Buckland and Elston (1993) suggest that even when many irrelevant variables are excluded from the analysis, the remaining variables may be highly correlated, leading to models which agree closely with observations but which give poor predictions when extrapolated to unsurveyed sites.

As a means of testing how these models for the alien study species compared to models for well-established species, logistic regressions were undertaken to predict the occurrence of well-established riparian species of a similar ecological niche. The results of such models using the Wear data (summarised in Table 4.29 below) showed similar levels of predictive ability to those of *Impatiens* and *Heracleum* within the catchment. This would tend to suggest that the ability to model these two aliens is no different from that of other species.

Table 4.29 Summary of predictive models produced to model the occurrence of *Phalaris arundinacea* and *Petasites hybridus* on the Wear

Species	% Correct			κ Values
	Present	Absent	Overall	
Phalaris arundinacea	73.66	73.98	73.98	0.48
Petasites hybridus	52.94	92.53	80.94	0.50

In order to alleviate the problems of creating a model based on one catchment alone, the amalgamated models were created. Such predictive models for *Impatiens* and *Heracleum* can be seen to put less emphasis on variables such as river width and seem to be more reliant on habitat variables. These models should be more applicable across a range of catchments, although several caveats still apply. Foremost amongst such caveats is the fact that a particular altitude, here probably acting as a proxy for climate, will not have the same climatic affect when applied to more distant catchments. Also the models do not take into account variation in the fine-scale differentiation of habitats, here classed into broad categories. For example deciduous woodland could encompass anything from wet alder carr to pioneer acid birch woods. An additional factor to bear in mind is the possibility of the species being able to occupy other habitats not available on the two study catchments. The only way to embrace such potential shortfalls is to include a wider range of river catchments in the data used for the creation of the model. However, even given these problems, it seems that reasonable predictive models for the occurrence of *Impatiens* and *Heracleum* can be created using crude habitat parameters.

Within catchment models produced to predict the occurrence of *Fallopia* are poor. This is attributed to *Fallopia* being dispersal limited due to its method of vegetative spread. Climate does not appear to be limiting this species in the catchments. This contrasts with the other study species whose seed dispersal mechanisms allow for rapid spread downstream and which are only really likely to be dispersal limited by long-distance upstream and inter-stream spread.

Many of the important variables in the single catchment and mixed-catchment models for the study species can be ecologically explained. For *Impatiens* the altitudinal limit has been noted elsewhere (Beerling & Perrins 1993) and has been suggested as representing a growing degree-day limitation (Beerling 1993). *Impatiens* is tolerant of a wide range of substrates and its negative associations with areas of bare silt and mud here are more likely to reflect its avoidance of the inter-tidal zones of the rivers than an actual substrate incompatibility. Similarly the positive association with increasing minimum river depth is likely to reflect the climatically suitable lowlands where the main rivers are generally deeper. A negative association with roads/bridges may be indicative of the higher frequency of bridges across the main river where it is narrower in the upper reaches of the catchment (see Figure 4.6), hence again reflecting climatic limits. Bridges also seem commoner in highly urban areas in the lowlands and may therefore also act as a proxy for unsuitable urban habitats. They are also frequent on the narrow tributaries and may reflect areas into which the species has yet to spread.

A negative relationship between the occurrence of *Impatiens* and short/mown grass and a positive relationship with unmanaged grassland relate to the species ability to colonise rank mesotrophic swards (Beerling & Perrins 1993) but its inability to tolerate prolonged grazing or other damaging management. Similarly positive relationships with herb/ruderal vegetation and woodlands are indicative of the species preference for tall herb-related communities (especially with species such as *Petasites hybridus* and *Myrrhis odorata*, Beerling & Perrins 1993) and the fairly open woodland canopies expected along the edge of rivers (Beerling & Perrins 1993).

Negative associations with areas classified as "wild" (undisturbed natural habitats) probably reflect the inability of the species to colonise well established, undisturbed natural vegetation types such as upland moors, upland and lowland grasslands and dense lowland woodland.

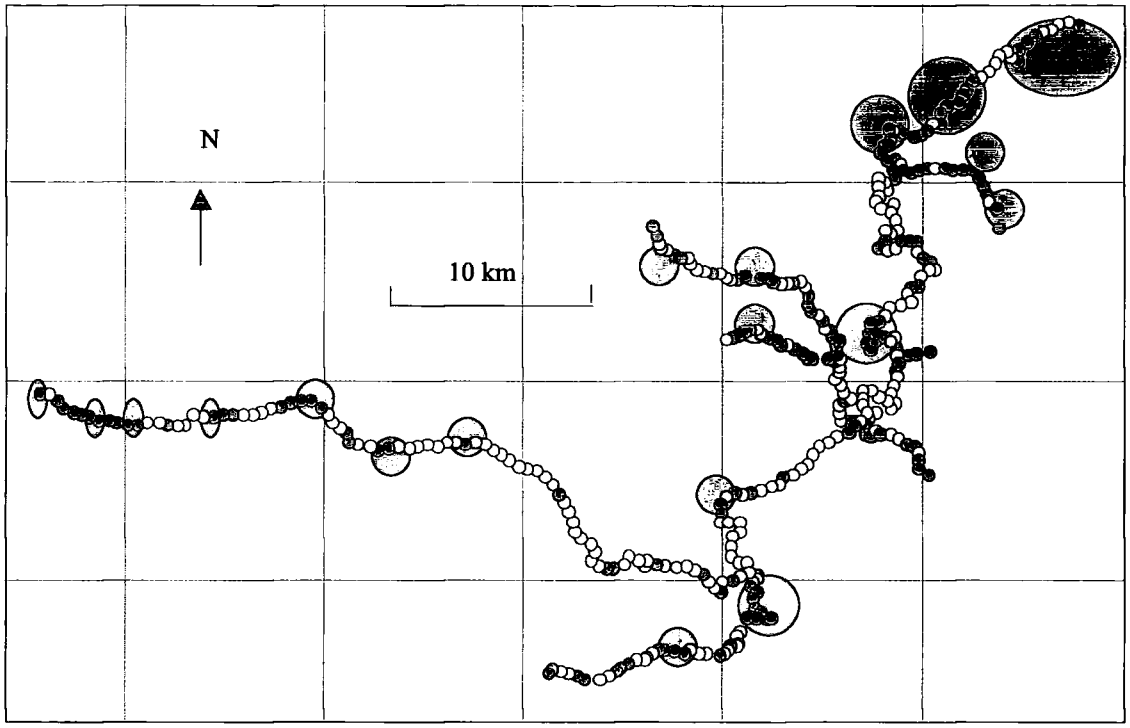

Figure 4.6 Distribution of the road/bridge variable along the Wear catchment. Urban areas along the river are shaded grey

As with *Impatiens*, so ecological explanations can be offered for the selection of most of the consistently included variables in the *Heracleum* models. The selection of altitude and maximum river width (both very much correlated) could reflect factors other than the most obvious explanation of climatic limitations. It is possible that they may be proxies for areas of slow water and potential seed deposition or perhaps of lowland floodplains with deeper soils. Published literature suggests *Heracleum* can occur up to about 200 metres in Scotland (Tiley *et al.* 1996), though on the continent it is recorded up to 1850m (Hegi 1905-31). The distribution through Britain and Ireland is currently predominantly lowland (Tiley *et al.* 1996) which corroborates the selection of altitude as an important variable in the models. Also, as *Heracleum* predominantly occurs in the main rivers of these catchments, a negative relationship with river width may reflect its inability to colonise the smaller tributaries.

The selection of woodland variables mirrors the species reported preference for woodland and grassland in other riparian areas (Gibson *et al.* 1995; though their "grassland" included species such as *Petasites hybridus, Urtica dioica* and *Symphytum officinale,* and would probably equate better to the RCS herb/ruderal variable). This preference for woodland may also reflect areas of deep soil or areas retaining soil

moisture, an important factor in seedling establishment (Neiland 1986, Tiley and Philp 1994). A noted preference for sheltered sites (Tiley *et al.* 1996) may also contribute to its association with woodland areas. The consistent selection of disturbed areas in the models (with a negative coefficient) is probably explained by *Heracleum*'s reported preference for undisturbed or unmanaged sites (Caffrey 1994). Neiland (1986) also reported that grazing and arable land were least often colonised. Disturbance also seems characteristic of the uplands of the catchment and of urban and inter-tidal areas in the lowlands (See Figure 4.7).

The inclusion of maximum bank slope and minimum bank height in some *Heracleum* models, both with negative coefficients may reflect a preference for low-lying areas of flat bankside. The inclusion of fine substrate as positive variables in the models (silt and mud) may be characteristic of particular zones of the river system or may reflect suitable seed depositional environments.

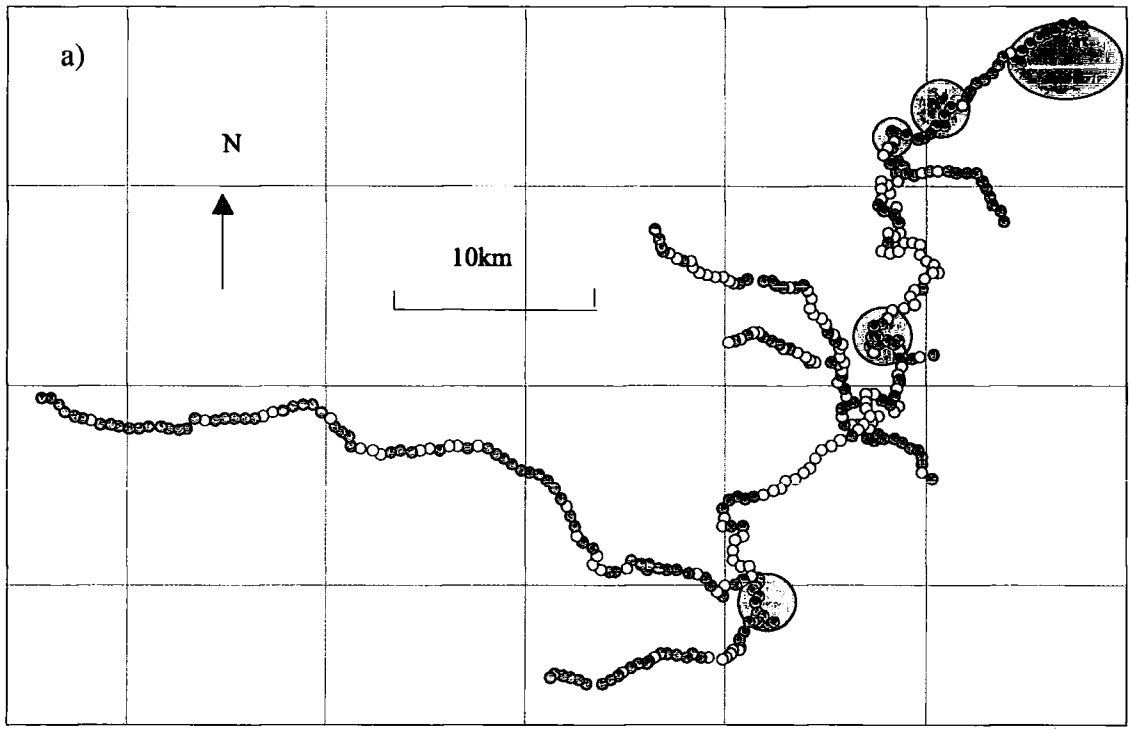

Fig 4.7 (see overleaf for title and legend)

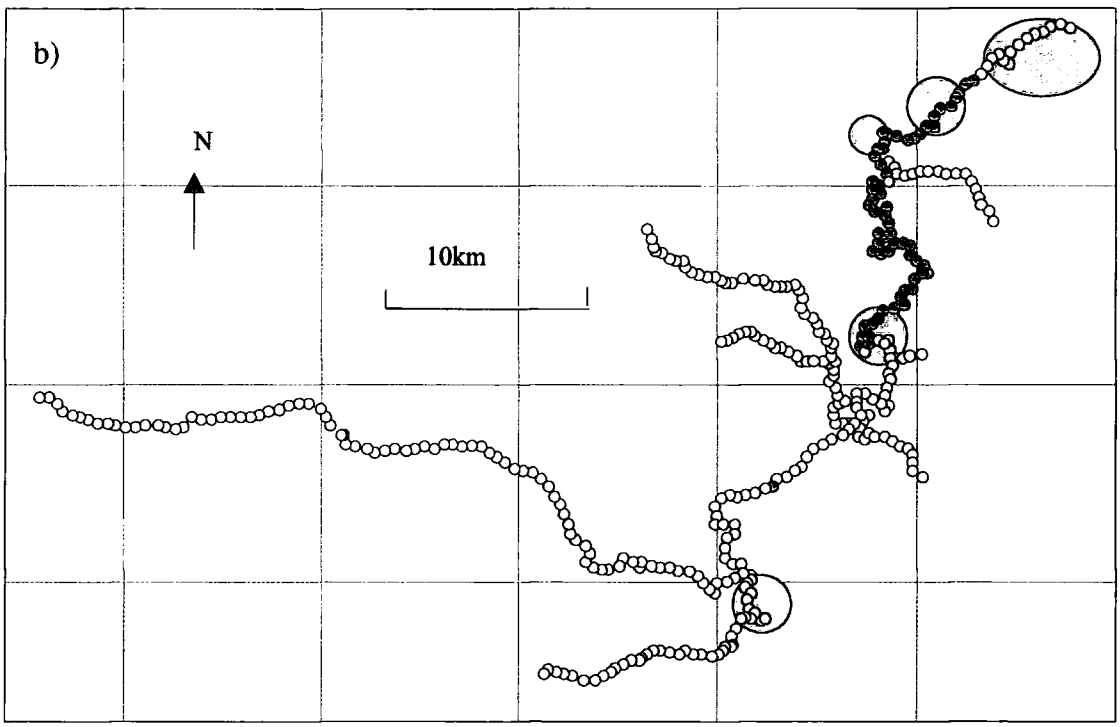

Figure 4.7 Occurrence of a) disturbed areas and b) *Heracleum* along the Wear. Major urban areas shaded grey.

The fact that within-catchment predictive models for *Fallopia* are poor is likely to indicate that the current distribution of this species is not limited by habitat or climate, though the latter has been shown to be a good predictor at larger scales (Beerling *et al.* 1995). If climate and habitat are not limiting distribution in the two catchments we may have to look to other factors such as dispersal limitations to explain the distribution of *Fallopia*. If the species is dispersal limited it is quite likely that it has yet to occupy its full habitat niche in the area. This may go a long way to explaining the inability to successfully model the distribution of this species at any of these levels. *Fallopia*'s late date of introduction to the two catchments, coupled with its reliance on vegetative spread for dispersal, aid in substantiating the theory that it is currently dispersal limited.

Ecological explanations can be offered for most of the common variables in the *Fallopia* models. The inclusion of deciduous woodland in several models reflects one of the species major habitat niches (Grime *et al.* (1988) indicate a preference for acid woodland), especially the open woodland fringes (Beerling *et al.* 1994) common along riverbanks. A negative association with areas of solid earth/rock cliff is likely to arise from the species intolerance of steep slopes (Grime *et al.* 1988). Positive associations with roads and residential areas/urban areas may reflect a combination of sites of introduction (garden escapes, dumped garden refuse, transported soil etc.) and suitable

sites of disturbance. Links with woodland and urban areas could possibly be related to a lower frost incidence (suggested as a potential limit by Beerling *et al.* 1994). Beerling & Palmer (1994) state that the species is less abundant above 150 metres in Wales, something also apparent from the Wear catchment but which cannot be accounted for by the simple presence/absence data of these analyses. They also found it to be absent above 460m, though on the Wear it does not occur above about 230m, again perhaps suggesting dispersal limitation.

The problem of trying to determine the mechanisms by which variables are operating on a species distribution is one that continually recurs in the catchment modelling of this chapter. This can probably only be properly resolved by devising suitable test experiments. It may be that a more logical solution to the modelling of species occurrences would be to adopt a two stage modelling approach; the first stage defining a suitable climatic window in which the species can persist, and the second using riparian variables to predict the actual occupation within this zone. This could potentially lead to better fits of the models and also to the inclusion of meaningful ecological variables which are currently not being selected by dint of their occurring both within and without of a species climatic window.

Chapter Five
Fine Scale Catchment Modelling

5.1 Aims and Introduction

The analysis of the RCS data presented in the previous chapter is based on the standard 500 metre stretches of river used by these surveys. The 500m recording unit can be considered somewhat crude when trying to associate species distributions to bankside habitat types, especially when a species is distributed locally within a section. This may explain some of the poorer model fits of the previous chapter; problems arising due to the scale of recording. The work in this chapter is therefore designed to refine the predictive abilities shown in the previous chapter, both by reducing the size of areas predicted and by using more quantitative species recording. Additionally the work will give a more accurate picture of the current distribution and patterns in distribution of the species.

The Wear catchment was re-surveyed along its entire length but at a resolution of 50m rather than 500m. By recording at 50m resolution and also by re-recording bankside habitat types it was possible to get a more accurate idea of the distribution of the three species in relation to riparian habitat. By additionally recording the study species' abundance and density at sites, a more robust dataset was available for analysis.

The aims of work presented in this chapter were to determine whether the occurrence of the three species, in terms of both presence and abundance, could be modelled at this finer resolution and whether such refined data would prove a better predictor of species occurrence. Resultant models may provide more of an insight into the ecological requirements of the species than the previous models and may shed light upon the individual species' meta-population dynamics.

Initial descriptive results presented examine how plant density and abundance change in different bank height zones. Models are then created using these data and their predictive ability examined at this finer scale.

Refining the RCS data

Plant species modelling at a catchment level is relatively unstudied as discussed earlier (section 4.1). This statement is truer still at the scale of the work presented in this chapter. Table 5.1 shows some of the major recent works undertaken on the occurrence and spread of the study species and gives an idea of the scale and extent of previous

studies. From this it is apparent that much previous work has completely overlooked habitat effects on the distribution of the species. Those studies including habitat effects have generally used very coarse, large-scale data. The collection of data at the 50m resolution adopted for this component of analysis falls somewhat between the two standard scales of ecological analyses, namely that of small scale studies, working with quadrats at a scale of metres, and that of larger scale analyses, generally working at resolutions of a kilometre or more. In deciding a scale at which to operate, both logistic and ecological considerations must be addressed. The disciplines of pattern analysis (Kershaw 1957, 1960; Greig-Smith 1957, 1961) and more recently metapopulation dynamics (e.g. Harrision 1994) both contribute to deciding the scale at which vegetation should be studied. Most studies of communities along riparian habitats work at the scale of 1-10 square metres (e.g. Rodwell 1998, 1999). The larger end of this range would be most suitable for the study species, being comparable to clump size for all three species. However clumps are often at an even larger scale than this, sometimes running over tens or hundreds of metres (e.g. Prach 1994; Tiley & Philp 1992). Fifty metres was chosen as the minimum size of surveying unit that was feasible to record yet would still provide useful quantitative information. Of the previous work operating at a similar scale (Table 5.1), most simply record species distribution patterns, the one exception being Beerling's work (1991a), though the habitat data he used were very coarse.

Table 5.1 Summary of the major works examining the occurrence and spread of the study species in native habitats.

Citation	Year	Species	Scale	Extent	Notes
Tickner	unpublished	*Impatiens*	quadrats & experiments	County	Examining the competitive interactions of *Impatiens* with *Urtica dioica* using both field quadrats and manipulative experiments.
Grime et al.	1988	*Impatiens, Fallopia*	quadrats	Br. Isles	Looks at a species habitat preferences using quadrat data.
Pyšek & Pyšek	1995	*Heracleum*	quadrat	25km^2 (Czech)	Uses quadrats to look at the occurrences of *Heracleum* across the entire landscape.
Prach	1994	*Impatiens*	patches/quadrats	2 x 1km areas	Uses quadrats to look at the dynamics of *Impatiens* in a small riparian area.
Neiland et al.	1987	*Heracleum*	patches	catchment	Fine scale distribution mapping along rivers but with no analysis of habitat selection.
Gunn	1986	*Impatiens*	patches	catchment	Fine scale distribution patterns of *Impatiens* along rivers.
Child & de Waal	1997	*Fallopia*	Patches	catchment	Using GIS and fine patch distribution to examine spread and control in the Swansea area.
Clegg & Grace	1974	*Heracleum*	patches	catchment	Fine scale distribution mapping
Beerling	1991a	*Fallopia*	200m	catchment	Examines the occurrence of *Fallopia* in relation to dominant land-use.
Hugh & David	1998	All three species	500m	Br. Isles	Use data from the River Habitat Survey to examine occurrence in relation to riparian characteristics.
Pyšek	1991	*Heracleum*	tetrad	Czech Rep.	Looks at large scale spread and frequency distribution in the whole landscape.
Conolley	1977	*Fallopia*	tetrad	Br. Isles	Solely examines spreading rates
Trewick & Wade	1986	*Impatiens*	tetrad	Br. Isles	Spread of *Impatiens* independent of habitat
Perrins et al.	1993	*Impatiens*	vice-county	Br. Isles	Spread of *Impatiens*, differing seed production
Beerling	1993	*Impatiens, Fallopia*	~50km^2	Europe	Uses climate alone to map the occurrence and potential spread of the two species.

5.2 Method

5.2.1 Defining the recording units for fine scale modelling

To define the 50m recording units the 500m RCS sections were first split into left and right banks. Each 500m of bank was then split into ten sections of approximately 50 metres by dividing the original RCS habitat map into ten approximately equal sections thus creating twenty new survey stretches within an original RCS survey stretch. Angold *et al.* (1996) in examining the accuracy of similar RCS surveys found that spatial location of RCS features could vary in the order of 2-25% of the total reach length. As most of the RCS data used in these analyses were based on large-scale maps (with exaggerated widths), section maps were usually found to tally quite well with distances on the ground and on OS maps. As the data collected were to be relative to the defined 50m sections, these locational errors were not considered important in this work.

As the RCS maps were drawn approximately to scale the adopted methodology provided a simple means of locating 50m sections, RCS habitat features being used to locate the beginning and end of sections in the field. In each of these new 50m sections the riverbank was split up into three different zones (Figure 5.1): the lower-riparian zone; the upper-riparian zone; and the non-riparian zone.

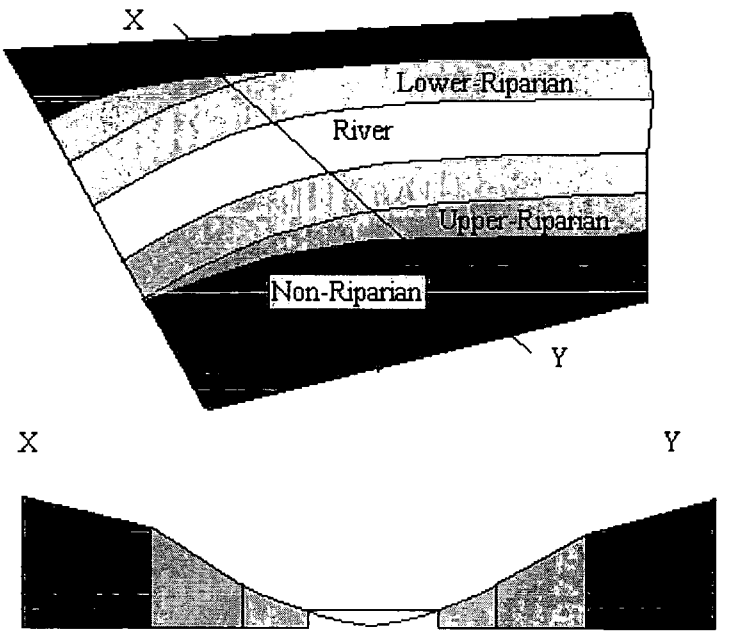

Figure 5.1 Stylised diagram and cross-section of the locations of the three zones of the riverbank in relation to the river itself.

To distinguish the three zones standard criteria were used. The lower-riparian zone was defined as that area below the height of the highest recorded strand line left by floods (presumably of the previous winter). If no obvious strand line was visible then this boundary was defined by the presence of other features indicative of flooded areas i.e. bare or sparsely vegetated areas, or the area where the dominant macrophyte species were those characteristic of ruderal/disturbed or regularly inundated areas. However in the field the latter methodology was rarely called into use, as there was usually a visible strand line that was found to tally very well with the limit of other ruderal signs.

The next zone up the bank from the lower-riparian zone was the upper-riparian zone. This zone, whilst usually being quite easy to distinguish at its lower limit (i.e. the strand line), was somewhat more arbitrary in the definition of its upper limit. In order to attempt to standardise its definition a list was drawn up of macrophytes which were identified as being characteristic of this zone (see Table 5.2). Although several of these species occurred in the lower-riparian zone their upper bankside limit helped to define the upper margin of the upper-riparian zone. This zone was thought to comprise a zone of the river that was only very irregularly flooded or which was under some of the localised effects of the proximate river (e.g. water level, slight climatic amelioration). It was occasionally the case that this middle zone was not present in the field, the bank grading straight from the lower-riparian areas into non-riparian habitat. This tended to occur in areas such as woodland or where the bankside was grazed. The non-riparian zone differed from the upper-riparian zone in containing only species characteristic of the local non-riparian landscape. The upper limit of the non-riparian zone could therefore not be distinguished from the wider landscape. As with the upper-riparian zone, the non-riparian zone was also occasionally absent when the river-bank immediately gave way to other land-uses such as arable farmland.

Table 5.2 Species characteristic of the upper-riparian bank habitat

Character Species	Character Species
Alnus glutinosa	*Petasites hybridus*
Arctium sp.	*Phalaris arundinacea*
Artemisia vulgaris	*Rorippa sylvestris*
Chamerion angustifolium	*Salix cinerea*
Epilobium hirsutum	*Salix fragilis*
Filipendula ulmaria	*Symphytum officinale*
Impatiens glandulifera	*Tanacetum vulgare*
Myrrhis odorata	*Urtica dioica*

For each 50m section, and within each of the three defined zones, the major vegetation and substrate types present within the section were recorded as present or absent in relation to a pre-defined list (see Table 5.3).

Table 5.3 Vegetation and substrate categories recorded during the fine scale studies

Vegetation categories	Substrate categories
reeds	silt
tall grass	mud
short grass	sand
mown grass	gravel
herb/ruderal	shingle
sparse herb/ruderal	cobbles
mixed grass/herbs	boulders
sparse grass/herbs	bedrock
sparse scrub	
dense scrub	
woodland	
occasional trees	

The presence of the three study species was recorded separately for the three zones in each recording section. To facilitate rapid surveying of sections simple DAFOR categories were utilised to record species abundance using the criteria shown in Table 5.4. The average density of the study species was also categorised for each zone in which they were present (Table 5.5).

Table 5.4 DAFOR abundance recording methodology for the three species

Abundance	Species distribution criteria
Rare	A few scattered plants or small patches
Occasional	occasional plants along the bank or several small patches
Frequent	distributed thinly along much of the length or several large patches
Abundant	well distributed along much of the length or numerous large patches
Dominant	dominant along large sections of the bank

Table 5.5 Density recording categories for the three species

Density Categories (plants/metre)	
0.5	(any < 1)
1	25
2	50
5	100
10	200

The resultant data, along with the general vegetation data for each stretch, could be used to examine species occurrence in relation to habitat types. This was expected to assist in the refining of previous predictive models and in bringing to light further factors (or reinforcing already important factors) influencing the distribution of the three species. Such information could be masked to some extent by the crudeness of the 500 metre data.

Data from this survey have the potential to be used in related catchment modelling (e.g. Collingham *et al.* in press; Collingham *et al.* 1997; Wadsworth *et al.* in press; Wadsworth *et al.*, submitted) to estimate potential carrying capacities. They may also prove to be of much use in the creation of refined statistical models that are able to predict more than just presence/absence of a species.

5.2.2 Timing and logistics of the survey

There was some discrepancy in the best time to record the three different species. All three species are at their most visible when flowering, which occurs at different times of the year for each species. As a compromise the months for surveying were chosen to run from June through to September/October. This also minimised the recording of very high spring germination densities.

Ideally when surveying the sections it would have been preferable to walk both banks individually, recording habitats and species after walking the entire section. However because of problems of access this was not always possible and it was sometimes necessary to survey both sides of a section from one bank. This method was found to be almost as accurate as surveying from both banks if a pair of binoculars was used and the far bank frequently scanned whilst walking the section. There were occasional problems in such surveying due to dense vegetation but generally the method seemed to work well.

5.2.3 Descriptive statistics

Initially the proportion of 50m sections recorded in any RCS section was examined to see how often and whereabouts the two upper bank zones were absent or not recorded. Following this the proportion of 50m sections that held each of the study species was plotted to see how the occurrence of the three species changed with distance upstream across the three bank height categories.

Using the data collected on the abundance and density of the study species simple exploratory statistics were used to look for similarities and differences in species occurrences.

5.2.4 Modelling using the refined distribution data and the RCS variables

As with the 500m RCS data, so it was necessary to validate the collected fine-scale data before analysis was undertaken. Autocorrelative checks were ran on the datasets and the data manipulated as necessary to reduce any effects.

Initially the 50m data were used as a means of quantifying species occurrences in the 500m RCS sections. By calculating proportional occupation of 50m stretches in a 500m section a quantitative value could be derived for occupation of any of the three bank height zones. It was hoped that these quantitative values could then be used as dependent variables in multiple regression analysis with the original RCS variables as independent variables. Unfortunately the proportion-occupation data could not be normalised so multiple regression analysis was invalid. Median abundances in the 50m stretches of a 500m RCS section were calculated and these values used in multiple regression analysis with the RCS habitat variables as independent variables. Spearman-rank correlations were also undertaken between the proportional occurrence of each of the three species and the RCS variables to examine associations.

5.2.5 Modelling with the 50m distribution data and the 50m resolution habitat data

Logistic regression analysis was used to see if the habitat and species distribution data collected at the 50m resolution could be used to accurately predict species occurrence at this finer scale. The analyses were undertaken on the data collected along one bank of the River Wear. These data were split into the three bank height zones and independent analyses undertaken for each zone. The distribution data for the three species were used as dependent variables in the logistic regressions. The habitat data collected at the 50m resolution were used as the independent variables along with several variables extracted from the original RCS data. Variables in this latter category included altitude, maximum river width, minimum river depth and the surrounding land-use variables; all of which would be approximately the same for each 50m section in a 500m RCS stretch. Autocorrelation analyses were again used to determine the 50m sections that could be safely used in the analysis.

5.3 Results

5.3.1 Bank zone recording

Figure 5.2 highlights the frequency with which the upper-riparian zone was partially or completely absent. The pattern of absences is similar through the length of the river with 50% or more of the 50 metre zones often not containing this bank category. Absences of the non-riparian zone by contrast rarely exceed 50% of the 50m sections in any RCS reach and are concentrated in the upper half of the catchment. This is likely to be a result of the narrower banksides of the upper river grading straight into farmland or being curtailed by boundary walls or similar features. All of the further work involving proportional occurrences of species corrects for these absences.

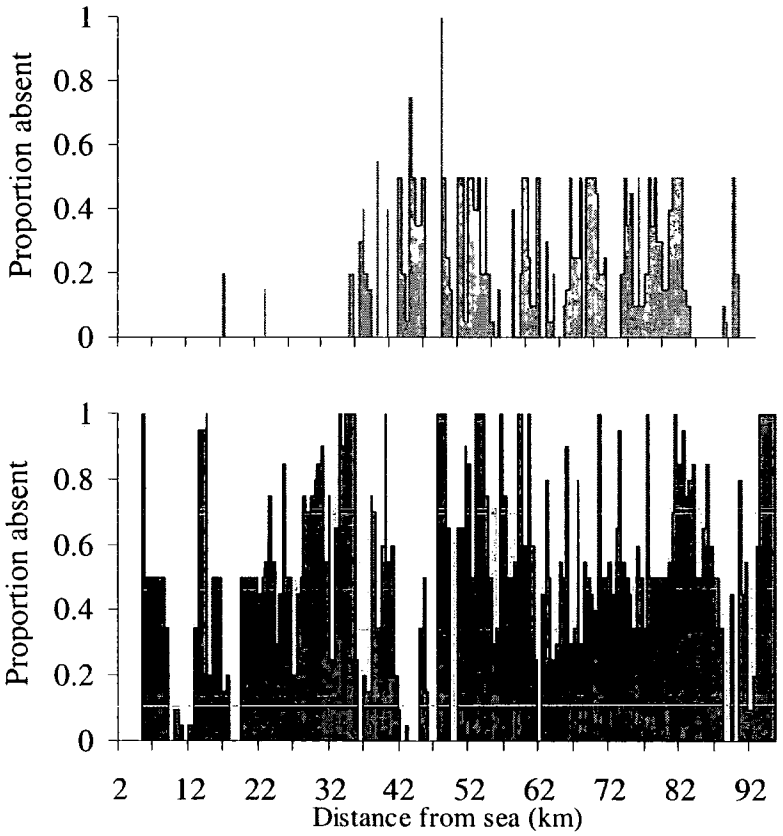

Figure 5.2 Proportional absences of the non-riparian zone (above) and the upper-riparian zone (below) along the River Wear. Absences are recorded as a proportion of 50m sections out of a 500m RCS section in which a zone did not occur. There were almost no sections in which a lower-riparian zone did not occur.

5.3.2 Species occurrence in the three bank zones

From Figure 5.3 it is apparent that the lower-riparian zone is by far the most important for the occurrence of *Impatiens*. The almost continuous distribution in the lower zone

contrasts sharply with a much more patchy and generally less frequent occurrence in the upper bank zones. Another notable fact is that the most downstream occurrences of *Impatiens* all occur in the non-riparian zone.

Unlike *Impatiens, Heracleum* rarely occurs throughout a 500m section in any bank height zone (Figure 5.4). However it is comparable to *Impatiens* in occurring most frequently in the lower-riparian zone, often occurring in 50-80% of the 50m cells in any section. Its distribution is patchy in the upper zones, generally occurring at lower frequencies where it does occur.

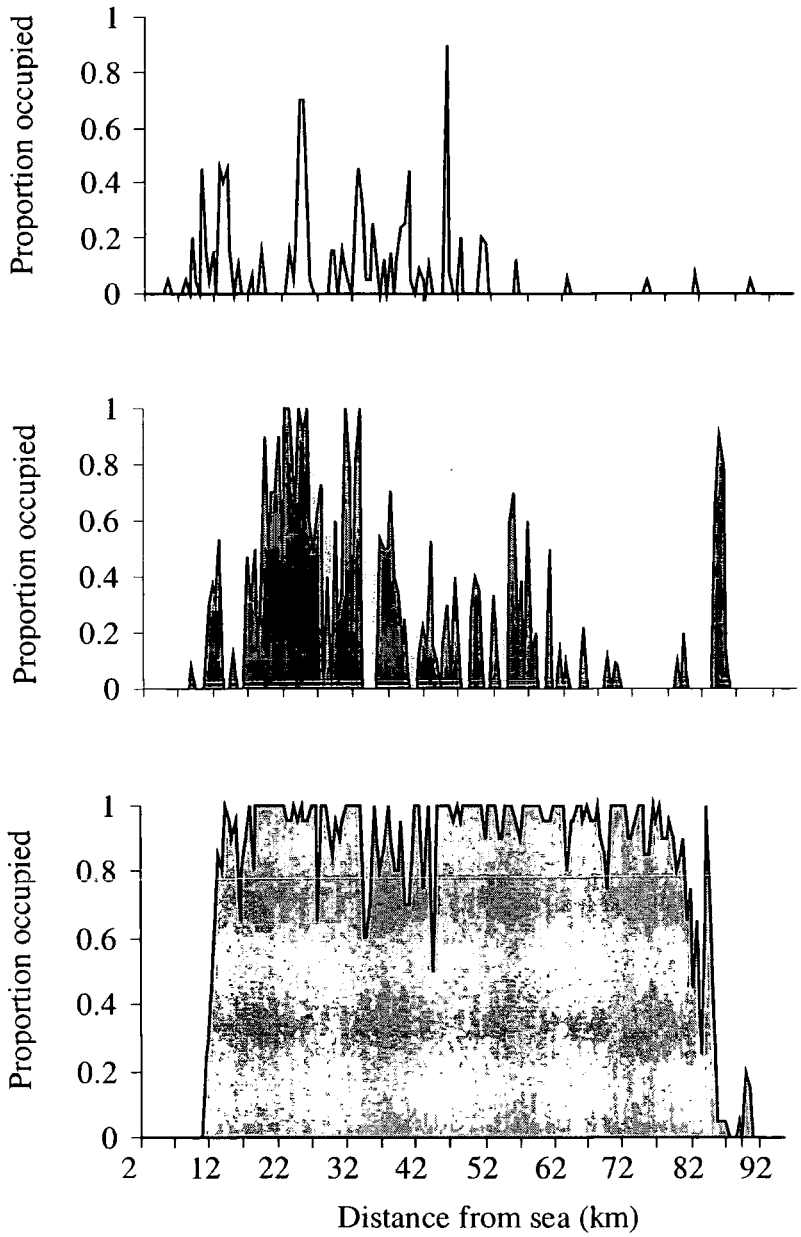

Figure 5.3 Charts of the proportion of 50m cells occupied by *Impatiens* in the 500m RCS sections of the River Wear. Proportional occupation is split into the three bank height zones of non-riparian (top), upper-riparian (middle) and lower-riparian (bottom).

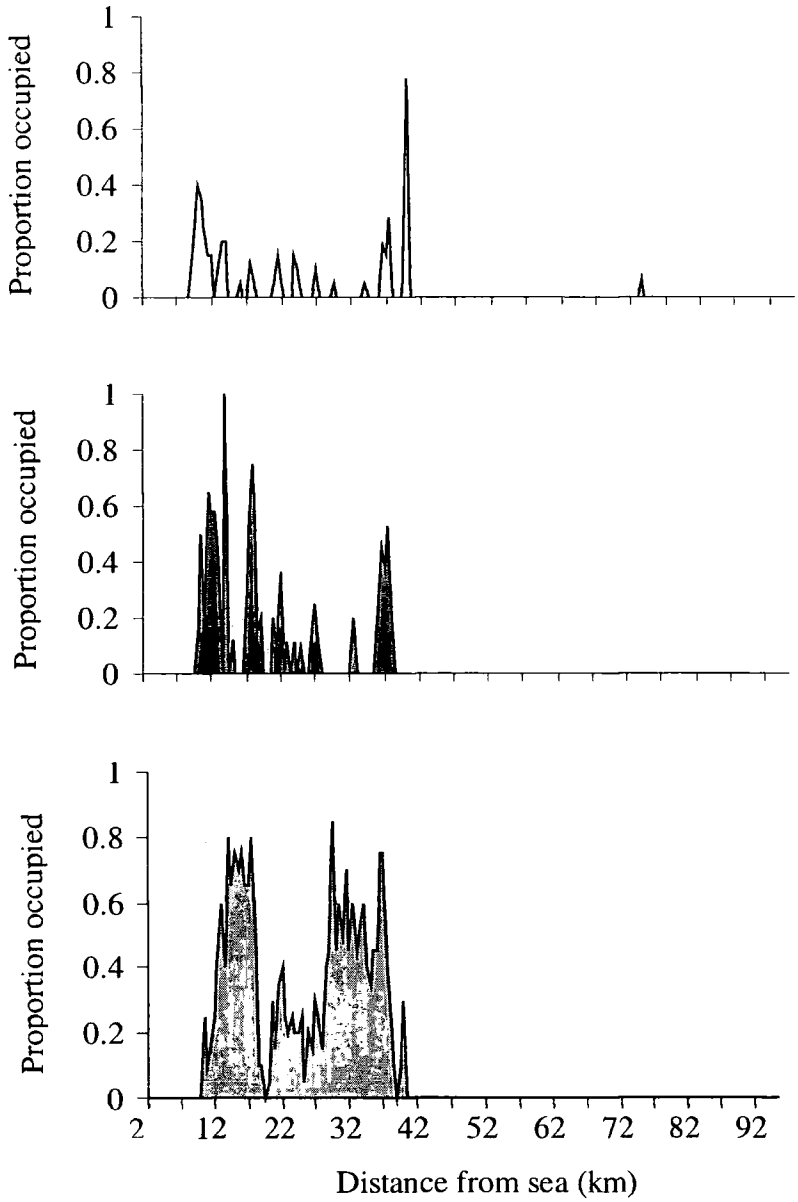

Figure 5.4 Charts of the proportion of 50m cells occupied by *Heracleum* in the 500m RCS sections of the River Wear. Proportional occupation is split into the three bank height zones of non-riparian (top), upper-riparian (middle) and lower-riparian (bottom).

The distribution of *Fallopia* (Figure 5.5) is much patchier than the other two species in the lower-riparian zone, only occupying a small proportion of 50m cells wherever it occurs. Occurrences in the upper-bank zones are infrequent and also indicate low occupation.

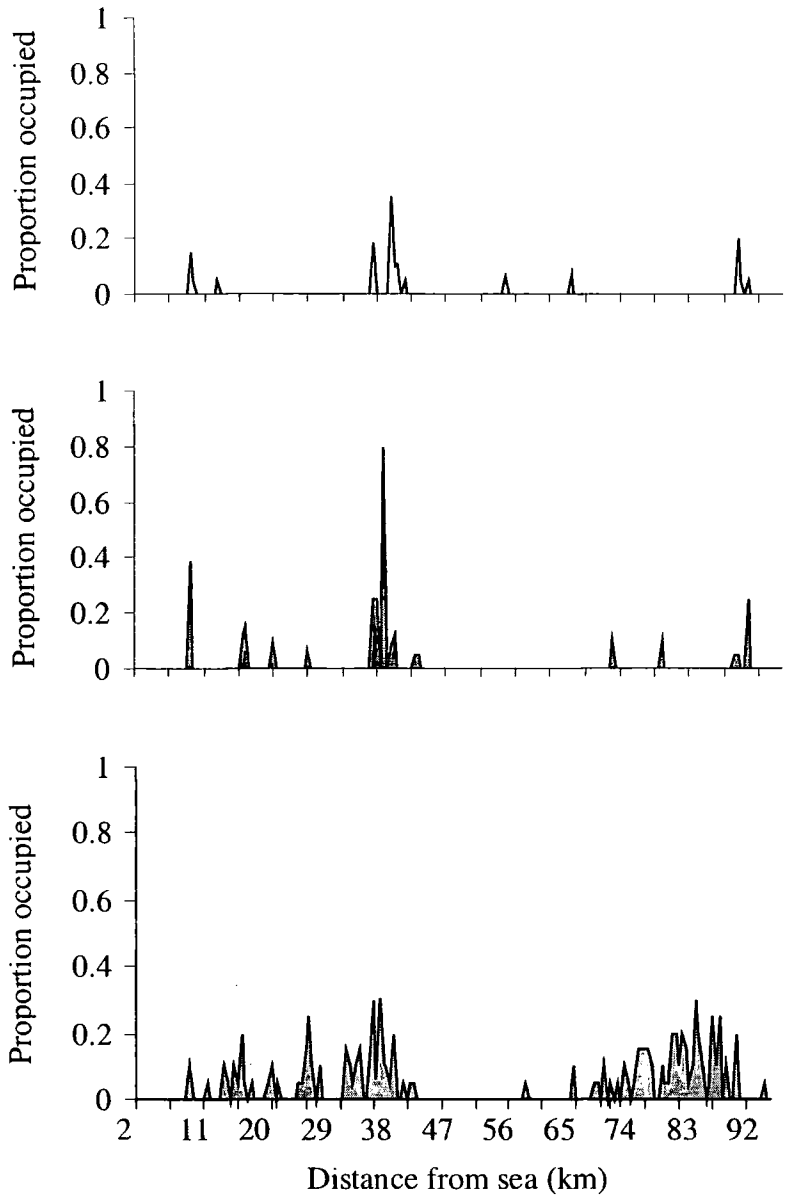

Figure 5.5 Charts of the proportion of 50m cells occupied by *Fallopia* in the 500m RCS sections of the River Wear. Proportional occupation is split into the three bank height zones of non-riparian (top), upper-riparian (middle) and lower-riparian (bottom).

Figure 5.6 summarises the recorded occurrences of the study species in the three bank height zones on the Wear. This highlights that the lower-riparian zone is the most frequently colonised zone for all three species. It must be borne in mind however that the upper- and non-riparian zones are often much reduced in extent or absent. *Fallopia* shows the most marked affinity to the lower-riparian zone, accounting for approximately 70% of its occurences. Imaptiens has approximately 65% of all its occurrences in this zone and *Heracleum* about 60% though the patterns of occupation of the zones by the three are not significantly different from each other. In terms of 500 metre reaches

containing the species, it therefore seems that the disturbed lower bank is by far the most important for harbouring populations of these species

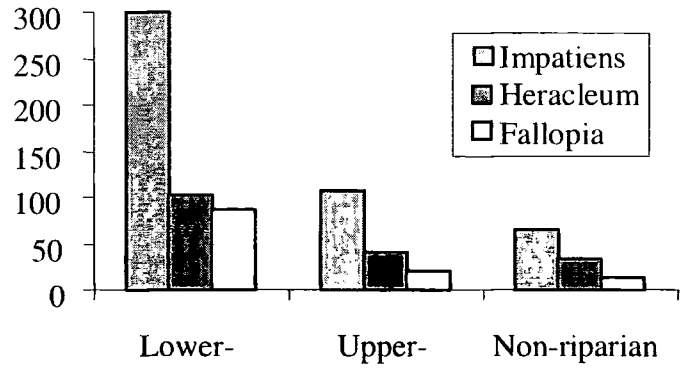

Figure 5.6 The number of occurrences of the study species in the 500 metre sections, split into the three bank-height zones along the River Wear. χ^2 =8.258; df =4; p = n.s.

The average number of 50 metre cells occupied in each occupied 500 metre cell is shown in Figure 5.7, summarising the data presented in Figures 5.3-5.5. From this it can be seen that if *Impatiens* occurs in a 500m reach, then it is likely to occur throughout that reach, especially in the lower zone. *Heracleum* also occurs most frequently in the lower zone but never as widely within a section as *Impatiens*. Conversely *Fallopia* shows no real difference in distribution across the three bank height zones, usually occurring in only one or two 50m zones in any occupied 500m section. These differences in distribution patterns between the species and also between the bank zones are statistically significant.

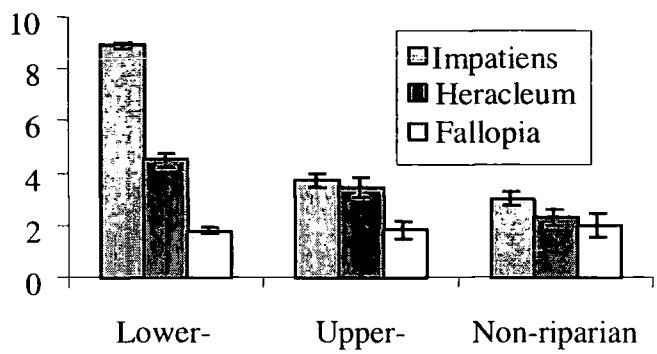

Figure 5.7 The mean number of 50m sections occupied in any 500m section in which a study species is present. ANOVA: Species; $F_{3256,2}$=83.31, P<0.01. Zone; $F_{3256,2}$=71.15, P<0.01.

5.3.3 Abundance distribution

The records of abundance of *Impatiens* (Figure 5.8) show a normal-type distribution with a marginal bias towards the lower abundances (rare and occasional). The species is generally recorded as being occasional to frequent and only rarely as dominant through an entire recording zone. This runs contrary to the often-perceived view of the species dominating whole areas of the riverbank. The trends are broadly similar for the three bank-height zones, though they are less obvious for the upper zones, which contribute fewer records.

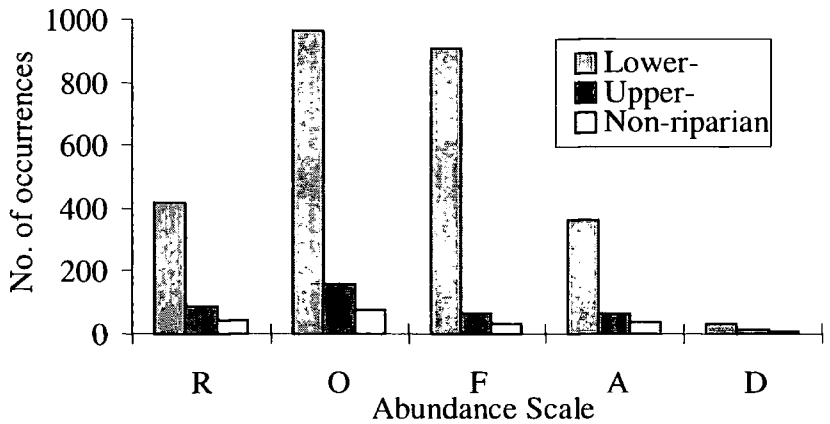

Figure 5.8 Frequency distribution of the recorded abundances of *Impatiens* along the River Wear, separated into the three bank-height zones. The categories of the abundance scale represent abundance as recorded using simple DAFOR criteria. χ^2 =80.27; df =8; P< 0.01.

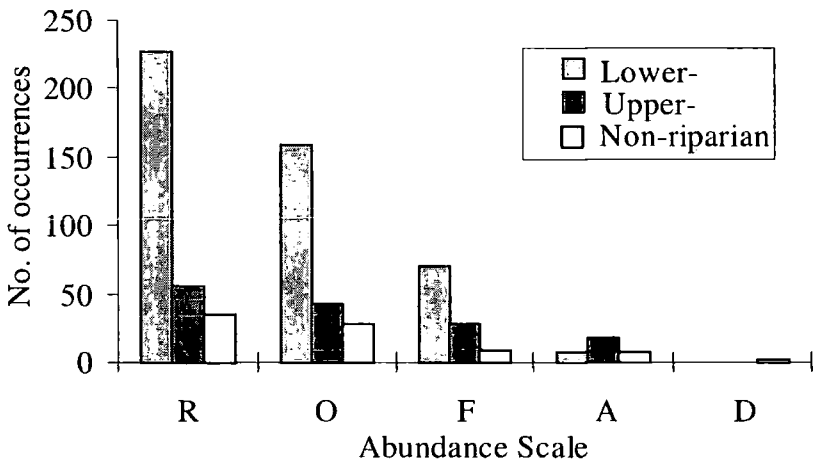

Figure 5.9 Frequency distribution of the recorded abundances of *Heracleum* along the River Wear, separated into the three bank height zones. The categories of the abundance scale represent abundance as recorded using simple DAFOR criteria. χ^2 =39.28; df =6; P<0.01.

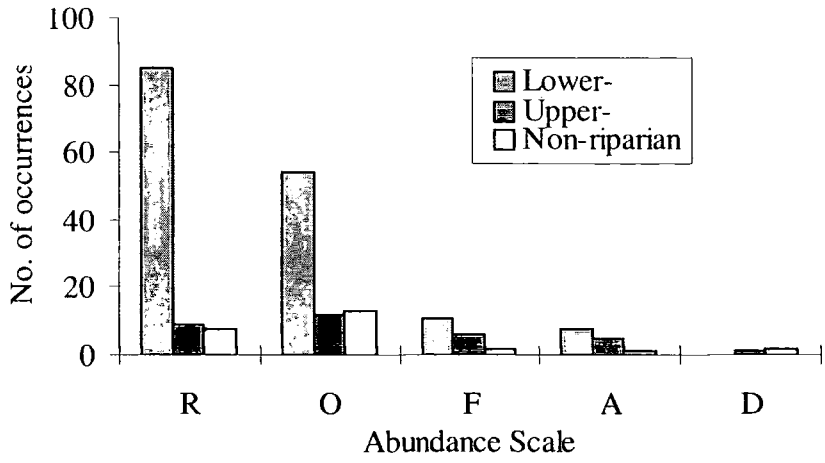

Figure 5.10 Frequency distribution of the recorded abundances of *Fallopia* along the River Wear, separated into the three bank height zones. The categories of the abundance scale represent abundance as recorded using simple DAFOR criteria. $\chi^2 = 17.76$; df =6; $P < 0.01$.

Records of abundance of *Fallopia* and *Heracleum* are very similar (Figures 5.9 & 5.10), both species occurring predominantly at low levels of abundance with fewer records as the abundance scale increases. Interestingly there are no records of either of these two species occurring as dominants in the lower-riparian zone, all such records are restricted to the upper zones.

5.3.4 Density distribution

Figure 5.11 shows the density distribution of *Impatiens* plants in the three bank height zones along the Wear study catchment. There is a normal-type distribution of densities in this ruderal zone with the species only rarely recorded at the highest and lowest densities, mostly being recorded at densities of 1-10 plants/m. The small number of high densities partly reflects the timing of surveying so as to miss the period of initial germination and early competitive exclusion. In the two higher bank zones the densities are more uniformly spread (except for the extremes). Apart from this fact, which may be an artefact of sample size in the upper zones, the data suggest no difference in density distribution of *Impatiens* between the three bank height zones.

The density distribution of *Heracleum* plants along the Wear has only a narrow range (Figure 5.12). Because of the narrow range of recorded densities no real trends can be detected in the species density distribution except to say that in all zones it is very rare at densities of more than 1 plant per metre.

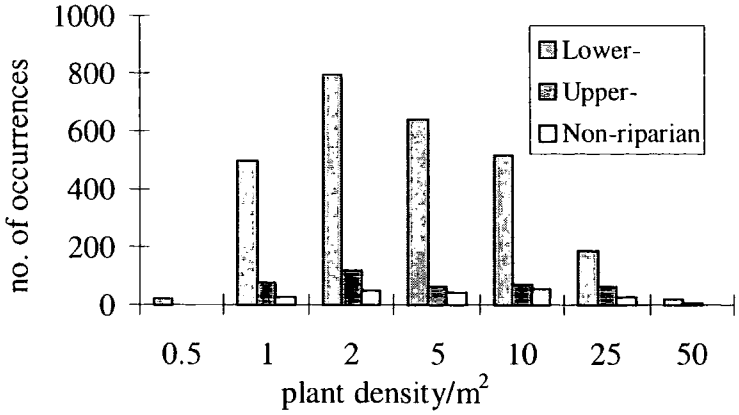

Figure 5.11 Frequency distribution of the recorded densities of *Impatiens* along the River Wear separated into the three bank height zones. χ^2 =21.31;df =12; P< 0.05.

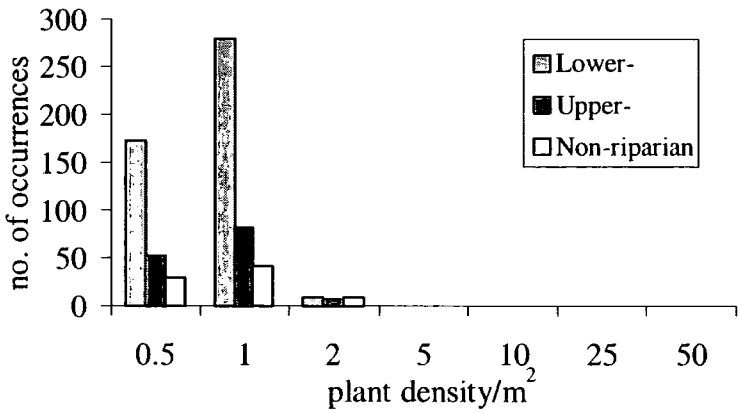

Figure 5.12 Frequency distribution of the recorded densities of *Heracleum* along the River Wear separated into the three bank height zones. χ^2 =14.71;df =4; P<0.01.

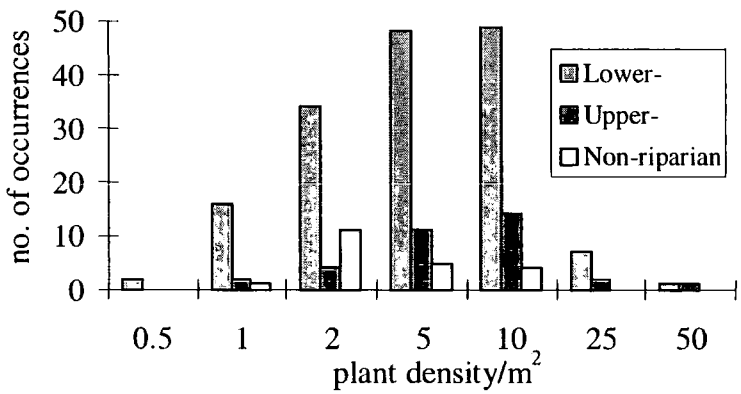

Figure 5.13 Frequency distribution of the recorded densities of *Fallopia* along the River Wear separated into the three bank height zones. χ^2 =13.24;df =8; n.s.

The density range of *Fallopia* (Figure 5.13) is much more skewed than that of the previous two species and unlike *Impatiens*, which is skewed slightly towards lower densities, the density range of *Fallopia* shows a decided skew towards higher within-

stand densities. This is most notable in the ruderal zone, less so in the upper-riparian zone and is not apparent in the non-riparian zone.

No obvious trends in differing stand density of each species between the three bank-zones are apparent, . There are differences in density distribution between the three species however, with *Heracleum* predominantly occurring at very low plant densities, *Fallopia* at high densities and *Impatiens* intermediate between the two

5.3.5 Using the refined distribution data and the RCS variables

Because of the large number of missing records in the upper-riparian zone the following analyses were undertaken using only the lower- and non-riparian zones.

Before undertaking any analysis using the small-scale data and bank height zonation it was necessary to check for associations both between the zones on a bank and between the same zone on opposite river banks. The data used in these correlation checks were the proportional occupation of 50m cells in a 500m RCS section. The results, displayed in Table 5.6, indicate significant correlations between the proportional occupation of a zone on one bank with that of the same zone on the opposite bank. The correlations between lower-riparian zones are much greater than between the non-riparian zones.

The results of correlating occurrence in different bank height zones on the same bank were less conclusive with only *Heracleum* showing a consistent correlation between zones and with no correlation coefficients being very large. As a result of these checks it was decided to keep the bank height zones separate but to combine the left and right bank proportional occupation data.

Table 5.6 Spearman-rank correlations between: a) proportional species occurrence in bank height zones on opposite banks of the River Wear; b) proportional species occurrence in different bank height zones on the same bank of the river Wear.

a)

Species	Zone	N	Spearman-rank	Sig.
Impatiens	Lower	181	0.639	0.01
Impatiens	Non-rip	181	0.284	0.01
Heracleum	Lower	181	0.828	0.01
Heracleum	Non-rip	181	0.359	0.01
Fallopia	Lower	181	0.206	0.01
Fallopia	Non-rip	181	-0.34	NS

b)

Species	Bank	N	Spearman-rank	Sig.
Impatiens	Left	181	-0.087	NS
Impatiens	Right	181	0.036	NS
Heracleum	Left	181	0.377	0.01
Heracleum	Right	181	0.382	0.01
Fallopia	Left	181	0.196	0.01
Fallopia	Right	181	0.098	NS

Partial-autocorrelation analyses, undertaken on the combined-bank proportional occupation data (Table 5.7) for the Wear, show that autocorrelative effects for all three species very quickly become negligible. This contrast with the strong autocorrelation detected in the simple presence/absence data (see Figure 3.1). For all three species autocorrelation between the neighbouring sections is always greatest in the lower riparian zone. These lower-bank autocorrelations are very strong for *Impatiens* and *Heracleum* (0.926 and 0.891 respectively).

Table 5.7 Partial autocorrelation figures of the proportional occurrence of the species in the three bank height zones along the 500m RCS sections.

	Fallopia		*Heracleum*		*Impatiens*	
Distance (km)	Lower-	Non-	Lower-	Non-	Lower-	Non-
0.5	**0.421**	**0.205**	**0.891**	**0.377**	**0.926**	**0.465**
1	**0.222**	0.102	**0.241**	0.077	0.140	0.012
1.5	0.132	-0.036	-0.108	-0.005	0.081	0.013
2	0.107	0.013	0.037	0.048	0.057	0.039
2.5	0.019	**0.222**	0.046	**0.250**	-0.067	0.041
3	0.098	-0.063	-0.004	0.010	-0.015	-0.080
3.5	0.111	-0.004	-0.114	0.096	0.014	0.083
4	-0.068	-0.010	-0.086	-0.077	-0.026	0.018
4.5	-0.028	-0.010	**0.151**	-0.067	-0.101	-0.024
5	0.097	-0.096	-0.113	-0.106	-0.021	-0.119
5.5	0.094	-0.013	-0.01	-0.016	0.036	0.108
6	-0.032	-0.03	0.058	-0.057	-0.066	-0.053
6.5	-0.084	-0.022	0.065	-0.027	0.129	0.024
7	-0.048	-0.023	0.055	0.021	-0.042	0.016
7.5	0.006	0.007	-0.07	0.078	-0.041	0.037
8	-0.056	-0.020	**-0.186**	0.038	0.082	0.107

Occurrence is recorded as the proportion of 50m cells occupied in any section along both river banks. Distribution data is from the River Wear. Figures in bold are those which lay outside of the 95% confidence limits (±0.146).

Spearman-rank correlations between the proportional occurrence of the three species and the RCS variables are shown in Table 5.8 for the two bank height categories.

Table 5.8 Spearman-rank correlations of proportional occurrence of study species with RCS variables in 500m RCS sections.

	Fallopia		*Heracleum*		*Impatiens*	
Zone	Lower	None	Lower	None	Lower	None
Gr/he					0.37***	
Altitude			-0.582***	-0.396***	-0.272***	-0.431***
Arable					0.432***	
Artific				0.177**	-0.343***	
Bed	-0.152*		-0.206**	-0.19**	-0.31***	-0.159*
Bould		-0.194**	-0.423***	-0.263***		-0.223***
Bridges		0.171*		0.181**	-0.243***	0.14*
Cobbles				-0.146*		
Conif_Ba						0.254***
Decid_Ba	0.153*		0.148*			
Decid_Lu			0.258***			0.308***
Max_dep	-0.135*				0.258***	
Min_dep					0.247***	
Derelict			-0.291***	-0.201**	-0.163*	-0.226***
Disturb			-0.171*		-0.411***	
Fal_lower		0.14*	0.228***			
Fal_non	0.14*			0.278***		0.233***
Farming	-0.142**	-0.222***				
Grazed			0.165*		0.137*	
Grass		-0.16*				
Hard_Cls			0.195**			0.207**
Her_lower	0.228***			0.496***	0.236***	0.506***
Her_non		0.278***	0.496***			0.391***
Human		0.142*		0.194**	-0.166*	
Imp_lower			0.236***			
Imp_non		0.233***	0.506***	0.391***		
Islands			0.231***			
Max_ht						0.136*
Min_slop		0.154*	0.395***	0.204**		0.263***
Mixed			0.29***	0.203**		0.177**
Mud			0.413***	0.317***		0.25***
Reeds			0.229***		0.189**	
Road_br			0.197**	0.228***	-0.209**	
Roads			0.287***	0.353***		0.195**
Scrub_ba		-0.138*				
Shingle			0.247***		-0.216**	
Sh_mo			0.181**			
Silt		0.144*	-0.208**		0.297***	
Sof_Slum	-0.261***					-0.166*
Stream				-0.154*		
He/ru			0.292***	0.138*	0.545***	0.197**
Tot_Res					-0.19**	
Tot_Wood						0.148*
Urban					-0.327***	
Weir					-0.238***	-0.169*
Max_wid	-0.144*		0.345***	0.264***	-0.349***	0.26***
Min_wid	-0.138*		0.382***	0.271***	-0.274***	0.296***
Wild	0.165*	-0.138*		0.191**		

Variable labels are as in Appendix I, Table AI1. Lower and None indicate the lower- and non-riparian bank height zones. Significance levels of 0.05, 0.01 and 0.001 are shown as *, ** and *** respectively.

Fallopia in the lower riparian zone shows a strong negative correlation with areas of soft/slumping banks and a positive correlation with *Heracleum* in the lower-riparian zone. However, in the non-riparian areas soft/slumping banks are not correlated with *Fallopia*'s occurrence, instead the occurrence of *Heracleum* and *Impatiens* growing in the non-riparian zone are positively correlated with the species, whilst farming in the surrounding landscape is a negative correlate.

In contrast to *Fallopia*, *Heracleum*'s occurrence in the two bank-height zones has many correlations in common. Negatively correlated variables in both zones include altitude, boulders, bedrock and derelict areas. Positive variables common to *Heracleum* in both zones include *Fallopia* (in the respective zone), *Heracleum* (in the alternative zone), *Impatiens* (in the non-riparian zone), minimum bank slope, mixed wood, mud, roads and bridges and river width. Variables that are only correlated in the lower-riparian zone include silt (negative), shingle, sort/mown grass, reeds, islands, *Impatiens* (in the lower riparian zone), hard earth/rock cliffs and deciduous woodland (land-use). Those only correlated in the non-riparian area include artificial banks, bridges and the human-influence variable.

Impatiens also has several variables strongly correlated with it at both bank heights. These include altitude (negative), *Heracleum* (in the lower riparian zone) and herb/ruderal vegetation. River width is strongly correlated with *Impatiens* at both heights but surprisingly has different coefficient signs between the two zones. Variables that are only correlated with *Impatiens* on the lower-bank include urban areas, residential areas (negative), silt (negative), shingle, reeds, disturbance (negative), river depth, artificial banks (negative), arable land-use and grass/herb vegetation. Variables only correlated with the species in the non-riparian areas include boulders (negative), bankside conifers, deciduous woodland (land-use), *Fallopia* (in the non-riparian zone), hard earth/rock cliffs, *Heracleum* (in the non-riparian zone), bank-slope, mixed woods, mud and roads.

Correlations of substrates with each of the species generally only occur in lower-riparian zone, suggesting habitat preferences for riverbank substrates, though negative associations in the non-riparian zone with boulders and positive associations with mud show that substrate may also be important on the upper bank. Several of the bank features are related to species occurrences on only one of the two height zones.

Examples are artificial banks, reeds and solid earth/rock cliffs with *Heracleum* and *Impatiens*, islands with *Heracleum* and soft/slumping bank with *Fallopia*. Other variables are equally correlated with occurrence of a species in both bank zones, e.g. areas of dereliction for both *Heracleum* and *Impatiens,* and minimum slope and the road/bridge variables for *Heracleum* only. Other variables, which are correlated with a species distribution in both zones, are of greater importance in one zone than another. For example the occurrence of herb/ruderal vegetation is much more important for the occurrence of *Heracleum* and *Impatiens* along the lower bank than the non-riparian areas, and weirs are more correlated with the occurrence of *Impatiens* along the lower bank.

It was assumed that the surrounding land-use variables should aid only the predictive ability of species in the non-riparian zone. The correlation results however show that this is not always the case. Deciduous woodland land-use, though correlated with *Impatiens* in the non-riparian zone is also correlated with *Heracleum* in the lower riparian zone. Similarly arable land is negatively correlated with *Impatiens* in the lower-riparian zone and farmland is negatively correlated with *Fallopia* in both zones, though the strongest correlation is in the non-riparian zone.

5.3.6 Modelling using species abundance

As a means of modelling species occurrences using the abundance values recorded at the 50m resolution, median values of abundance were calculated for each of the RCS sections along the Wear. These could then be used as dependent variables in multiple regression modelling with the correlated RCS variables as the independent variables. Unfortunately only *Impatiens* in the lower-riparian zone created enough non-zero median values to justify undertaking such analysis. The result of this multiple regression modelling is displayed in Table 5.9 and Figure 5.14. These show that a relatively good-fitting predictive model can be created incorporating species' abundances and using only a few variables.

Table 5.9 Summary of the fit of the model created to predict the proportional occurrence of *Impatiens* in the lower zone of the River Wear.

a) Model Summary

R	R Square	Adjusted R Square	Std. Error of Estimate
0.796	0.634	0.625	0.774

b) Coefficients

Variables entered	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
constant	2.750	0.222	-	12.36	0.000
altitude	-0.009	0.001	-0.602	-12.57	0.000
wid_max	-0.302	0.030	-0.458	-10.20	0.000
arable	0.624	0.125	0.229	4.99	0.000
reed	0.085	0.026	0.152	3.34	0.001
cobble	0.424	0.169	0.112	2.51	0.013

Section a) summarises the fit of the observed against the modelled predictions and section b) shows the variables selected in the model and their coefficients.

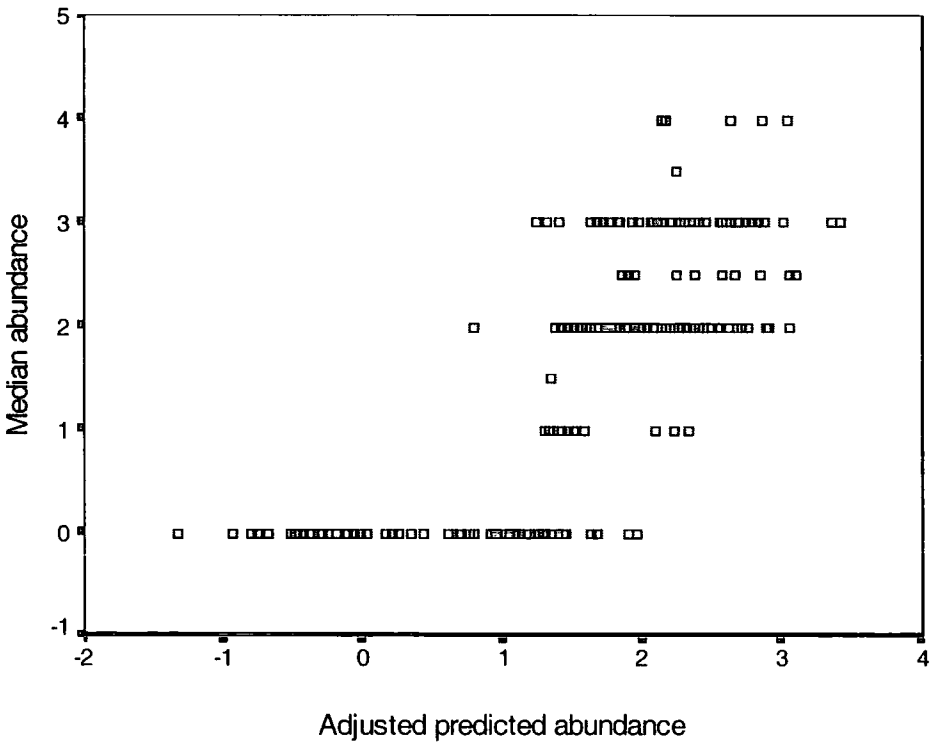

Figure 5.14 Median abundance of *Impatiens* in the lower riparian zone of 500m RCS sections against abundance as predicted using multiple regression with the RCS variables.

5.3.7 Analyses using the 50m resolution data of both species and habitat type

In these analyses the upper-riparian zone was not omitted, as there were enough occurrences of it at this 50m resolution to warrant its retention. The table of autocorrelation between the 50m sections for each of the three species (Table 5.10) indicates definite differences in autocorrelation between zones for any one species and also differences between species. The distance downstream for which autocorrelation is significant is taken as the point at which the first non-significant autocorrelation occurs. Hence the occurrence of *Fallopia* clumps are consistently autocorrelated only for 100-150m in all three bank-height zones. *Heracleum* occurrence is autocorrelated for 350m in the lower bank zone but only for 250m in the upper-riparian zone and 100m in the non-riparian zone. The distribution of *Impatiens* follows a similar trend to *Heracleum*, being autocorrelated to 600m in the lower bank zone but only 200 and 250m in the upper zones. The autocorrelation values are far higher in the neighbouring 50m section than any other sections suggesting that autocorrelation falls off rapidly at first and then tails off more slowly.

Table 5.10 Partial autocorrelation figures of the presence of the three species along the 50m sections of the River Wear.

	Fallopia			*Heracleum*			*Impatiens*		
Distance (m)	Lower-	Upper-	Non-	Lower-	Upper-	Non-	Lower-	Upper-	Non-
50	**0.256**	**0.355**	**0.458**	**0.554**	**0.545**	**0.515**	**0.802**	**0.597**	**0.522**
100	**0.058**	**0.165**	**0.118**	**0.328**	**0.28**	**0.162**	**0.272**	**0.178**	**0.21**
150	0.023	**0.056**	**-0.037**	**0.184**	**0.148**	-0.031	**0.165**	**0.136**	**0.101**
200	**0.055**	0.02	-0.017	**0.105**	**0.037**	-0.003	**0.183**	**0.066**	**0.134**
250	0.06	0.02	0	**0.139**	**0.083**	**0.097**	**0.101**	0.07	**0.053**
300	-0.001	-0.026	-0.024	**0.066**	0.017	0.004	**0.086**	**0.037**	0.01
350	0.029	0.012	0.015	**0.082**	**0.051**	0.026	**0.063**	**0.069**	-0.018
400	0.026	**-0.045**	0.008	0.1	**0.075**	**0.042**	**0.086**	**0.054**	**0.069**
450	0.022	0.024	**0.059**	0.025	**0.035**	**0.054**	**0.071**	0.007	-0.019
500	-0.03	-0.018	-0.03	0.04	**0.088**	**0.045**	**0.079**	**0.038**	0.021
550	**0.044**	**0.062**	**-0.044**	**0.135**	0.005	0.015	**0.034**	0.009	-0.003
600	0.027	**0.059**	0.01	0.019	0.006	0.05	**0.057**	-0.027	0.027
650	0	-0.003	0.008	**0.078**	**0.051**	**0.115**	0.024	0.008	-0.001
700	0.028	0.002	**0.048**	**0.039**	0.019	**-0.091**	**0.047**	0.007	-0.027
750	0.005	0.048	0.003	0.109	0.023	0.011	0.012	0.047	0.024
800	-0.012	0.021	0	-0.016	-0.073	0.001	0.061	0.027	-0.025

Figures in bold are those which lay outside of the 95% confidence limits (±0.034).

To take into account these effects of autocorrelation, the 50m sections used to create regression models were chosen so as to negate these effects. Therefore for *Impatiens* in the lower riparian zone only every thirteenth section was used in the analysis, whereas in the upper riparian zone every fifth section was used. Similarly for *Heracleum* in the lower riparian zone only every eighth section was used. However when this method was adopted for the species in the other bank height zones it was found that there were not enough occurrences of the species to produce reliable models. Therefore for all analyses, bar the three mentioned above, all of the occurrences of a species in a zone were used in the model so long as they did not occur within the distance of autocorrelative effects of another site. An equal number of non-occurrence sites were also randomly chosen and these two datasets used to create the logistic regressions. In the case of *Fallopia* in the two upper-bank zones there were not enough sites of occurrence to justify undertaking any analyses.

The results of these analyses are given in Table 5.11 and show a mixture of predictive abilities varying from excellent to poor. Predictive ability can be seen to vary considerably between the different bank height zones, as can the variables used to create the models for any species (Table 5.12).

Table 5.11 Summary of the regression models using species presence/absence data at 50m resolution with the 50m habitat data and selected RCS variables.

Species	Zone	n	Chi2	df	Classification P	A	total	κ
Impatiens	Lower-	110	101.241	6	100	100	100	1
Impatiens	Upper-	156	35.285	5	37.5	95.69	80.77	0.40
Impatiens	Non-rip	117	66.424	6	87.93	76.27	82.05	0.64
Heracleum	Lower-	203	62.139	3	41.94	97.09	88.67	0.47
Heracleum	Upper-	51	42.919	2	92	88.46	90.2	0.80
Heracleum	Non-rip	90	83.245	7	93.33	91.11	92.22	0.84
Fallopia	Lower-	87	18.005	4	73.33	57.14	65.52	0.31

Table 5.12 Variables used in the creation of the 50m logistic regression models summarised in Table 5.11.

Varaibles	*Impatiens*	*Impatiens*	*Impatiens*	*Heracleum*	*Heracleum*	*Heracleum*	*Fallopia*
Model	Lower-	Upper-	Non-rip	Lower-	Upper-	Non-rip	Lower-
Constant		---	---	NS	+++	NS	--
He/ru	+++	++	++			+++	
Sp_he/ru							+
Decid_ba			+++	+			++
Occ_tree	+++						
Scrub							+
Gr/he	+++		+		-	+++	
Mown		NS					
Mud							+
Sand		+					
Bed				++			
Altitude	---		-	---	--	--	
Arable		++					
Scrub_lu			-				
Conif_lu	---						
Mixed_lu						+	
Roads						++	
Min_dep		-	++			+	
Max_wid	---					-	

+ and – are used to denote significant positive and negative coefficients; multiples of 1,2 and 3 representing significance levels of 0.05, 0.01 and 0.001 respectively. Variables are as in Appendix I, Table AI1 with the addition of Sp_he/ru and Occ_tree, for sparse herb/ruderal and occasional trees respectively

The variables used in the two good *Impatiens* models (lower- and non-riparian zones) are quite similar; both selecting herb/ruderal vegetation, grass/herb vegetation and woodland categories as positive influencing variables and both with altitude as a negative influencing variable. Contrarily the *Heracleum* models have no variables in common except altitude, though most are ecologically feasible selections. Similarly for the *Fallopia* model, most variables selected are ecologically sensible, with the exception perhaps of scrub.

The RCS surrounding land-use variables were included in the models in case they were influential on species distributions in the non-riparian zone but, with the exception of the road and mixed wood variables used in the *Heracleum* model, this seemed not to be the case.

5.4 Discussion

The proportional occupancy results for the three species show three very different patterns. The almost continuous distribution and relative abundance of *Impatiens* and *Heracleum* in the lower-riparian zone could be expected as a result of these species' water-borne dispersal strategies. The extent to which *Impatiens* occupies all the lower-riparian areas in its range reflects the tenacity of the species to survive in even small areas of suitable habitat and must also be indicative of very good dispersal ability. Mechanistic models by Wadsworth *et al.* (submitted) substantiate this.

The spread of *Impatiens* into the non-riparian zone beside the river however is shown to be a far from regular event. Occurrence in the upper riparian zone is somewhat intermediate between that of the lower- and non-riparian zones. Only in the lower reaches of the river does it seem to be able to establish itself in the upper-riparian zone with any regularity.

Given the hydrochorous seed dispersal mechanism of *Heracleum* (i.e. floating on water) it may be expected that this species would occur with equal regularity in both the lower and the upper-riparian zone, seeds being deposited on the border of the two zones. This is obviously not the case and, as with *Impatiens*, the upper zones support much reduced numbers compared to the lower bank. The continual flooding of lower bank might be expected to be detrimental to a perennial such as *Heracleum* which could be washed out in floods but judging by its ability to occur widely in this zone this is not a problem.

Fallopia, spreading slowly by vegetative means or by fragmentation would be expected to have a patchier occurrence than the other two species and at lower frequency, which is borne out by the presented results. The predominance of occupation along the lower-bank could also be expected given the slow vegetative horizontal spread compared to a fairly rapid dispersal by water when fragmentation occurs. Of the few occurrences in the non-riparian environment, it is interesting that two such sites lie at the very upstream end of the major areas of occurrence. It is possible that these non-riparian introductions facilitated vegetative spreading to the lower-riparian zone and then a rapid subsequent hydrochorous downstream dispersal, causing the present lower-riparian distribution.

The results of abundance recording seem to disprove the often-stated fact that these invasive species dominate whole areas of riverbank to the exclusion of all other species (Beerling *et al.* 1994, Beerling & Perrins 1993, Tiley *et al.* 1996). At the scale of 50m recording the species rarely ever dominate the bankside. This may be due in part to the large scale of the recording unit or perhaps the non-equilibrium distribution of the three species at present. It may be that abundance will increase further as the species consolidate their spread. *Heracleum* and *Fallopia* are most regularly recorded as rare or occasional, most of these records coming from the lower-riparian zone. The predominance of small patches of *Heracleum* and *Fallopia* has been noted elsewhere (Tiley *et al.* 1996, Beerling & Palmer 1994). The reduction in records at increased abundances is less marked in the upper two bank zones suggesting that when *Heracleum* and *Fallopia* occur on the upper bank, they are more likely to occur at higher levels of abundance than they are in the lower zone.

Impatiens is more frequently recorded at higher levels of abundance and the trends are similar across all three bank-zones. It is much less likely to be recorded as a rare component of the bankside vegetation than the other two species, predominantly occurring at an occasional or frequent level of abundance, and is quite often recorded as abundant in the lower-riparian zone.

The predominance of records of occurrence, at all levels of abundance, in the lower-riparian zone has important implications for the further spread of these species. By concentrating most individuals in this zone closest to the river the species are facilitating rapid range expansion by placing most of their propagules within easy reach of the river, encouraging long-distance hydrochorous dispersal. It seems a fortuitous pre-adaptation to prosper in such environments is facilitating their current rapid expansion.

The recorded densities in the field show very different patterns for the three species, rational ecological explanations accounting for much of this variation. *Impatiens* tends to occur most frequently at densities of 1-10 plants/m^2. These values seem low compared to values published elsewhere (Beerling & Perrins 1993 {50-70 plants/m^2}, Perrins *et al.* 1993 {30-40 plants/m^2}). However it must be remembered that the densities recorded here are an average over an entire zone, i.e. not choosing the most dense patches, and also that the surveying was carried out from June when intra-specific competition had already greatly reduced numbers. The values better match the recorded densities of Prach (1994), who observed initial densities of over 100 plants/m^2, falling to <10/m^2, with the greatest mortality occurring between May and June, and those of Prowse (1996).

Recorded densities of *Heracleum* compare very well to those noted by Tiley *et al.* (1994) who state that mature flowering specimens are commonly no more than 0.5-1/m^2. The densities of *Fallopia* shoots seem slightly low compared to other recorded densities in Britain (Beerling 1990a {~14/m^2}, Scott 1988 {~40/m^2}) but high compared to densities recorded in its native Japan (0.032-0.39/m^2, Maruta 1983). Recording average densities over all sites in a zone probably explains the slight discrepancy between these and the other densities recorded in Britain. The small clumps, only recently established by fragmentation, are also likely to contribute to the number of low-density records. The skew towards higher shoot densities in the two lower zones and not in the non-riparian zone is contrary to expectations. It would be expected that stands in the non-riparian zone, that can only spread vegetatively, would be denser on average than those in the lower zones, due to the continual establishment of small new colonies in these low-lying areas.

The rapid reduction in autocorrelation between RCS sections when quantitative values are applied is very different to the long autocorrelation tails presented previously using simple presence/absence data. It therefore seems that whilst the presence of a species in one section means it is likely to occur in adjacent sections, it is not possible to predict the abundance of a species in adjacent sections based on a known quantitative score. This is logical, as for example the mass input of propagules from an upstream colony may well lead to the establishment of a few plants in suitable areas downstream but there is no guarantee that the bankside in such a section will contain suitable habitat types for large scale colonisation.

The Spearman-rank correlations indicate that there are differences in the variables correlated with species distributions in the different bank height zones. The variables most highly correlated to species distributions in each of the zones are reassuringly those most pertinent to that zone. For example, by and large, when two of the invasive species are correlated it is the occurrences of the species in the same bank height zone that are correlated, suggesting similar habitat selection rather than correlation based on large-scale spatial niche similarities. Strong correlations occur between the occurrences of all three species, particularly in the non-riparian zone, suggesting some similarity in habitat selection.

From the correlation analyses it is possible to build up pictures of riverine habitat selection for the three species. *Impatiens* populations become less likely to occur at the upper end of the catchment, whereas at the seaward end they can occur only above the lower-riparian zone. Deep sections of river increase the liklihood of the species in the lower zone, whereas disturbance reduces its occurrence there. Areas of dereliction are generally avoided, especially so when they occur in the non-riparian zone. The occurrence of *Heracleum* increases the liklihood of the species occurring in both zones, the species being more likely to occur in the non-riparian zone when *Fallopia* is also present. The species is also more likely to occur in the non-riparian zones when much of the bank is steep, the surrounding land-use includes woodland or where there is an element of unnatural woodland along the bank. The presence of reeds and herb/ruderal vegetation along the lower bank often encourages the species whereas proximate weirs have the opposite effect.

Heracleum is most likely to occur in both zones in the lower, wider reaches of the river, avoiding areas of bedrock, boulders and derelict habitat wherever they occur. *Fallopia* and *Impatiens* occurring in a bank height zone increases the probability of *Heracleum* occurring, as does the occurrence of *Heracleum* in the adjacent bank height zones of a reach. Areas with much of the bank steep, containing mixed woodland and areas of bare mud encourage *Heracleum* in both zones, whereas the presence of reeds, shingle and herb/ruderal vegetation encourage it in the lower-riparian zone.

Fallopia remains difficult to predict, with fewer strong correlates than the other species. The best predictor in both zones seems to be the occurrence of *Heracleum*, though from the distribution maps *Fallopia* obviously occurs out of this species' range. *Fallopia* in the non-riparian zone is more likely to occur when *Impatiens* is present, whereas

farming practices reduce the likelihood of its occurrence. Similarly areas of soft earth cliffs and slumping bank reduce the chance of it occurring in the lower-riparian zone.

These correlations for the three species generally match the variables shown to be important in previous chapters and fit with the available literature of each species habitat associations (see the discussion in the previous chapter for details). Where it differs from the previous modelling is in showing that some of the variables only encourage the species in particular zones of the riverbank and not throughout. Variables such as herb/ruderal vegetation are shown to most strongly influence species occurrence only in the lower bank, whereas others such as derelict land have a similar effect wherever they occur.

The modelling of *Impatiens* in the lower zone using median values of abundance indicates that the production of models to predict abundance is possible, though some refined method of recorded abundance or an increase in records is necessary if the other species and other zones are to be modelled. The variables used in the model are similar to those in the *Impatiens* models of the previous chapter with the additional inclusion of reeds; highlighting the previously overlooked importance of this variable in the lower zone.

The autocorrelation analysis of the species at 50m resolution indicates differences in the dynamics of the three species that are not obvious in the 500m analyses. It can now be seen that the occurence of *Fallopia* is much less dependent than the other species on neighbouring cells being occupied and that autocorrelation is similar in the three bank zones. This fits with the low rate of spread by means of fragmentation and vegetative growth. The high levels of autocorrelation of *Impatiens* and *Heracleum* in the lower zone reflect the importance of long-distance hydrochorous spread for these species, *Impatiens* seemingly having the better dispersal mechanism. This fits in with models of *Heracleum* and *Impatiens* seed dispersal (Wadsworth *et al.* in prep.) that use biological and physical parameters of the two seeds to predict downstream dispersal distance.

The logistic regression models, produced using presence of the species at 50m resolution in the different bank height zones and associated habitat variables, are highly variable in their predictive abilities. The results don't support the idea that the species could be better-predicted in any bank height zone by reducing the scale of the recording units. Good predictive abilities for *Impatiens* in the lower-riparian zone contrast with

poor predictive powers for the other two species and good predictions for *Heracleum* in the upper zones are not matched by *Impatiens*.

The variables chosen in the models are also quite different between zones for any species. Herb/ruderal vegetation is very important for *Impatiens* in all zones, though only in the non-riparian zone for *Heracleum*. This latter fact may indicate that *Heracleum* requires periodic disturbance or open areas to establish in non-riparian areas. Though woodland is important for *Impatiens* in the non-riparian zone (similar to the previous modelling using proportional RCS occupation) it is the presence of occasional trees in the lower-riparian zone which best predicts *Impatiens*. The presence of grass/herb vegetation is most important in the lower-riparian zone, presumably where there are still enough open areas to allow *Impatiens* to grow without being out-competed. A negative relationship between *Impatiens* and river width occurs in the lower-riparian zone, indicating an avoidance of wide inter-tidal areas.

Variables in the upper-riparian *Impatiens* model are quite different to those in other models, which may explain its poor predictive ability. Similarly those variables in the lower-riparian *Fallopia* model are, with the exception of woodland and herb/ruderal vegetation difficult to explain ecologically and consequently lead to poor predictive ability.

The modelling at these smaller scales and using more quantitative data has yielded important differences between these and the models of the previous chapter. Differences have been shown in the choice of variables to predict occurrences in differing bank height zones. Predictions have been shown to be possible at a smaller scale (though with varying accuracy) and also to some extent with quantitative data. Significant differences in distribution, abundance and density patterns have also been shown for the three species, most of which can be rationally explained.

The vast improvements in predictive ability expected using the refined data were not realised, perhaps suggesting that larger-scale analyses with coarse data are the most appropriate methods for predictive analyses of these species. Conversely it could be suggested that higher definition data (both physical and ecological) are necessary for more accurate small-scale and quantitative predictions.

The high degree of autocorrelation for *Impatiens* and *Heracleum* and the more local scale of autocorrelation for *Fallopia* using 50m resolution data, can be compared to the

previous autocorrelation values (Chapter 4) using the 500m resolution data (>5km for all species) and other autocorrelation noted at larger scales. Collingham *et al.* (in press) found that spatial autocorrelation occurred in the distribution of *Fallopia* and *Impatiens* at the tetrad scale but not with *Heracleum*, whereas only *Impatiens* demonstrated significant spatial autocorrelation at a hectad scale.

The findings presented here, in light of the other published autocorrelation results highlight variables degrees of autocorrelation with scale, probably determined by species dispersal traits and niche width at the various scales. Both in the work presented here and also in work published by other members of the research group (Collingham *et al.*, in press) there is a trend of better fitting logistic models when a species occurs widely at the particular scale of study. This may well reflect the theory that good logistic models cannot be created at any scale when a species is not yet at equilibrium in its distribution.

From the results shown in this chapter we can infer that the three species, despite marked similarities in their successful spread as aliens, have very different population and meta-population dynamics. Levins (1970) in his initial definition of meta-populations referred to extinction-prone con-specific populations. The distribution changes shown earlier, at a 500m resolution, would indicate that extinctions do not occur in these systems. However field-work has shown that populations of both Heracleum and Impatiens are prone to local extinction, despite overall persistence in a 500m section, between one year and the next. It is therefore reasonable to consider the study species in a meta-population context.

Impatiens being an annual and well-dispersed species would be expected to be very variable in its distribution from year to year. Its occurrence in numerous patches along the river fits the classic patchy meta-population system, with local extinctions likely to be countered by seed influx from upstream populations. Its limited occurrence on the upper riverbanks suggests that these areas are acting as a propagule sink.

Heracleum, like *Impatiens*, is a well-dispersed species but unlike *Impatiens* it is much less widespread with only a few very large patches. Although the species shows no sign of extinctions between surveys at a 500m resolution, the frequent populations of just one or two plants suggest many short-lived small colonies that vary in their exact location over time. This is therefore analogous to a mainland-island meta-population structuring.

Fallopia fits the description of a species displaying non-equilibrium meta-populations. Its poor dispersal ability and the relative isolation of small populations means any local extinction is unlikely to result in re-establishment. However the occurrence of large founder populations at the upstream end of its two main population patches in non-riparian habitat, which are supplying propagules for dispersal along the lower-riparian bank, fits a mainland-island-type population structure.

In the respect of continual downstream dispersal, the populations of all three species could be compared to source-sink populations, with upstream sites acting as a constant propagule source.

Chapter Six

Species Performance in Riparian Habitats

6.1 Aims and Introduction

The work presented in this chapter examines the effects of riparian habitat-type within the species' ranges on the performance, and hence potential for spread, of the species. This work highlights variation in performance between habitats, which should aid the understanding of realised and future species spread. By comparing variation in habitat-related performance with the predictive models previously created it will be possible to see if the results of these demographic studies are reflected in the modelled occurrences of species.

Although the studies of habitat were undertaken on all three species, *Fallopia*, occurring as a clonal species, is not amenable to most demographic analyses in this chapter. For this reason much of the chapter deals with the two other species, for which this approach is more applicable.

A method was proposed by Crawley *et al.* (1993) for predicting the invasion of alien species by calculation of an invasion criteria ($\lambda_1>1$), taking account of a species demographic traits. The finite rate of increase can be most easily estimated by sowing a known number of seeds and monitoring their subsequent survivorship to adulthood and resultant fecundity. However the invasive nature of the study species precludes introducing these species into natural environments (it is an offence to do so for *H. mantegazzianum* and *F. japonica* under Schedule 9 of the Wildlife and Countryside Act 1981). An indirect approach adopted here estimates invasiveness by assessing the survivorship of individual plants from seedling to adult and measuring their seed production.

It is often cited that disturbance is a major factor in determining the ability of invasive species to colonise an area (e.g. Hobbs 1989). This assumption was examined in two ways; firstly by monitoring fecundity and survival characteristics in a frequently disturbed habitat compared with more stable habitats, and secondly by disturbing some riverbank areas and comparing plant performance in these to that of undisturbed control areas.

Published literature has provided data on performance of the species (Beerling 1990a; Salisbury 1961; Tiley *et al.* 1996) and their habitat (Merry *et al.* 1981; Usher 1986), community (Grime *et al.* 1988; Beerling 1990; Graham 1988; Gibson *et al* 1995) and substrate preferences (Beerling & Perrins 1993; Beerling, *et al.* 1994; Coombe 1956; Clegg & Grace 1974; Neiland 1986), but no work has so far linked the two fields of plant performance and habitat selection. Such data are essential if accurate predictive models of spread are to be constructed.

By examining species demography in light of invasion criteria theory, this chapter adopts a bottom-up approach to predicting invasion, which can be contrasted with the top-down approaches using habitat data from the RCS and smaller-scale surveys to predict occurrence.

6.2 Method

6.2.1 Experimental design

Studies were undertaken in areas representative of the typical broad habitats utilised by the three species. From the field surveys of the River Wear it was thought that the three habitats best representing the occurrences of the three species were similar and could be put into the following three categories:

1. Woodland/woodland edge communities (woodland in text)
2. Herb/ruderal or disturbed communities (ruderal in text)
3. Grass/herb or undisturbed open communities (grassland in text)

Though somewhat broad definitions, these three habitat types were thought to be adequate to measure any noticeable differences in demography, which in turn could reflect the distribution of the three species in the River Wear catchment. These habitat types were noted as the most abundant habitat types along the riparian systems, and also those habitats within which the three species most commonly occurred. The habitats are analogous to those used in the RCS mapping criteria. The ruderal vegetation is associated with the lower-riparian bank areas (as defined in chapter 5) whilst the other two habitat types are associated with the upper-riparian areas of bank.

6.2.2 Experimental protocol

For each of the selected habitat types three replicate areas were identified for each species, within which demographic studies could be undertaken. The replicate areas for all three species were all within a short distance of Durham City in the lowlands of the

county, an area in which the species thrive. These areas were within close proximity of each other, minimising any variation in demographic parameters due to local climate.

At each of the three sites for each species, five replicate plots measuring 1x1m were selected at random from both the periphery and centre of stands of each invasive species. In practice however suitable grassland sites containing *Heracleum* and *Fallopia* were difficult to locate in the study area. This habitat was therefore excluded from the *Fallopia* studies and, though some sites were found for *Heracleum*, these were made invalid for work other than germination studies as a result of subsequent bankside management.

In the autumn of 1996 half of each plot was cleared of all surface vegetation and cultivated to a depth of 15cm by hand digging (disturbance treatment) and half the area was left undisturbed. These plots were then monitored in each month during the growing season of 1997, recording natural regeneration of the study species and the subsequent survivorship of plants. Colour marking of emerging seedlings was used as a means of detecting different germinant cohorts. As well as recording monthly germination and subsequent survivorship, some measures of propagule output were also recorded at the end of each growing season. In the case of *Fallopia*, as propagule output could not be assessed and germination *per se* did not occur, the only measures that could be recorded were shoot numbers.

The original quadrats set up in the autumn of 1996 were, where possible, followed through the duration of the 1997 and 1998 summer seasons. Unfortunately many of the original quadrats were lost during the period of study due to a variety of factors, but mainly as a result of autumn/winter floods either washing away quadrat sites or burying them under alluvium. Because of this a proportion of quadrats had to be replaced at the start of each growing season, losing some of the desired continuity.

6.2.3 Analysis

From the survivorship data it was possible to examine the mortality of *Impatiens* and *Heracleum* between monthly survey periods to establish whether or not it was density dependent. This was undertaken using data over the duration of entire growing seasons and subsequently by splitting the data in monthly blocks.

Survivorship analyses of *Heracleum* and *Impatiens* plants growing in the permanent quadrats over the two growing seasons were undertaken using life tables. Several assumptions were made in undertaking these analyses. Firstly it was assumed that all

germination took place at one time and that there was no early or late germination. This assumption is valid for the studied populations as all seedlings germinated *en masse* with little late germination. A second assumption that had to be made was that any mortality that took place between recording intervals occurred mid-way between the two periods. As the collected data were censored i.e. not all plants were followed to mortality, Kaplan-Meier estimators (Norusis 1994) were used to calculate survival times and significance levels.

In addition to calculating growing-season survivorship it was also possible to calculate over-winter mortality for *Heracleum* seedlings. This was done by comparing the number of surviving seedlings in quadrats at the end of the first growing season, with the number of such plants growing in the following spring. Because *Heracleum* plants can flower any time from two years old onwards it was not possible to estimate overall mortality to maturity without following the quadrats over several more years.

An estimate of seed production of both *Impatiens* and *Heracleum* plants was made using data gathered from all three study-areas of each species. At each site of a study species the seed output of approximately fifteen plants was estimated each year. For both species the seed output data were derived from plants examined towards the end of their period of seed production (September to October). For *Impatiens* the total number of seedpods produced was calculated by totalling the number of pods on the plant, the number of sites of already spent pods and the number of flowers flowering at the time of examination. It was assumed that all the flowers at this late stage of the season would mature seed but that any subsequent flowers were likely to fail. To then derive seed estimates from these pod totals it was necessary to calculate a mean number of seeds per pod. This was achieved by the collection of between 10-20 pods on each plant, gathered from a range of terminal and axillary racemes, and at a range of distances along the raceme. A mean value of seeds per pod could then be calculated for each plant and hence a mean seed estimate derived.

To estimate seed output of mature *Heracleum* plants several plant characteristics were recorded for each individual. Initially the presence of a terminal umbel and the number of satellite and axillary branch umbels were recorded (see Tiley *et al.* 1996 for terminology). For the primary umbel the number of rays on the umbel was noted and, for a segment sample of 10 rays, the number of pedicels with flowers in these secondary umbels was also recorded. From this it was possible to estimate the seed output of the

terminal umbel, assuming that each flower produced 2 seeds, using the following simple equation:

Seeds on terminal umbel = 2(mean number of pedicels/ray x terminal ray count)

Similar calculations were undertaken using the satellite and axillary umbels (usually one of each) to derive average seed output per umbel and then multiplying by the total number of such umbels. Totalling all these umbel counts provides an estimated seed output per plant.

6.3 Results

6.3.1 Seedling numbers

The quadrat experiments highlighted variation in the propagule production of the three species both between years and also between habitats (Figures 6.1-6.3). *Heracleum* and *Impatiens* both show high germination in ruderal areas and lower germination in woodland, with an average of 200-500 seedlings/m^2 produced in ruderal areas and 50-100 seedlings/m^2 in woodland areas. By contrast *Fallopia* produced only low shoot numbers, though also demonstrated a trend of increased shoots/m^2 in ruderal sites.

Habitat type significantly affects numbers of both *Heracleum* and *Impatiens* seedlings, being higher in ruderal habitats than in woodland and grassland areas. Although there are differences in seedling numbers between years for both these species the trends are not significant. 1998 was, on average, a poorer year for germination of both *Impatiens* and *Heracleum*, (with the notable exception of *Impatiens* in ruderal areas) but was a better year for shoot production by *Fallopia* in the two study habitats, the latter reflected in a significant between-year effect. The large increase in shoots in ruderal areas in 1998 is biased due to new quadrats happening to contain high densities of shoots. A comparison using only quadrats surveyed over both years temper this large mean to a level of about 18 shoots/m^2, other values remaining similar. Nevertheless these annual differences remain highly significant. The effects of habitat-type on Fallopia shoots are also highly significant, as is the habitat-year interaction.

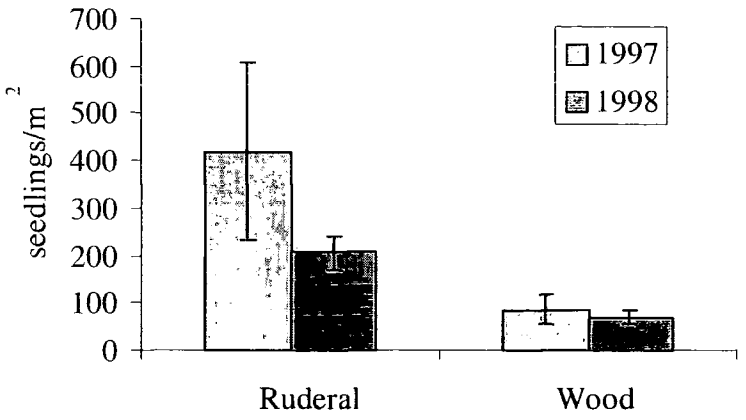

Figure 6.1 Germination of *Heracleum* seedlings in the habitat quadrats. Data are shown for spring of both 1997 and 1998. ANOVA: Habitat ($F_{1,50}$=10.23) P<0.01; Year ($F_{1,50}$=2.49) NS; Habitat x Year ($F_{1,50}$=1.81) NS.

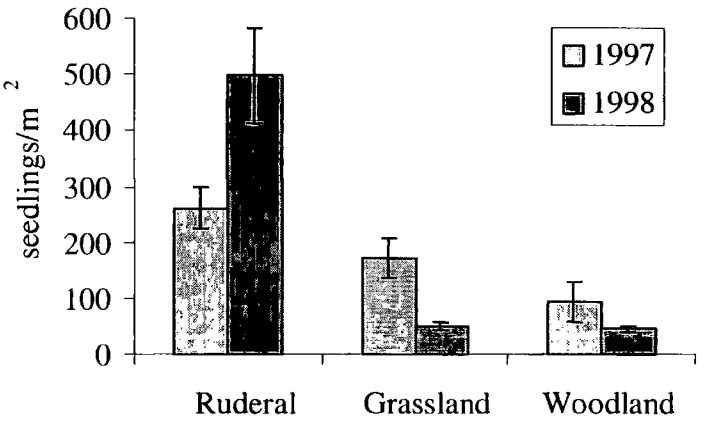

Figure 6.2 Germination of *Impatiens* seedlings in the habitat quadrats. Data are shown for spring of both 1997 and 1998. ANOVA: Habitat ($F_{2,86}$=42.09) P<0.001; Year ($F_{1,87}$=0.576) NS; Habitat x Year ($F_{1,87}$=12.986) P<0.001.

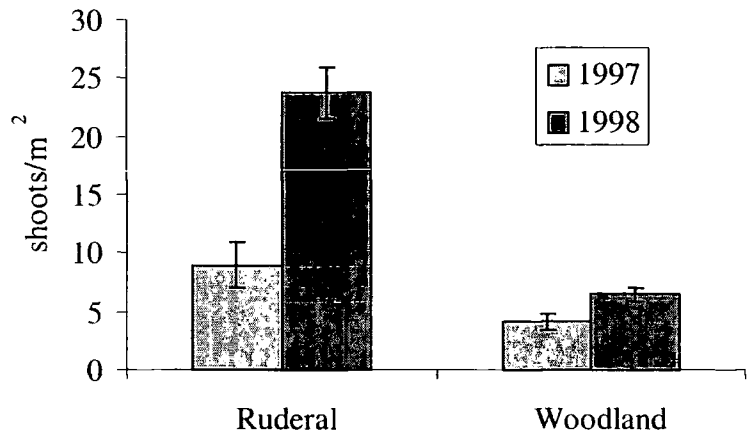

Figure 6.3 Maximum number of *Fallopia* shoots in the habitat quadrats. Data are shown for summer of both 1997 and 1998. ANOVA: Habitat ($F_{1,64}$=51.62) P<0.001; Year ($F_{1,64}$=30.70) P<0.001; Habitat x Year ($F_{1,64}$=16.517) P<0.001.

6.3.2 Disturbance treatment

The effects of disturbance on *Heracleum* seedlings and their subsequent survival are shown in Figures 6.4 with paired t-tests or Wilcoxon's signed rank tests (for data with unequal variance) used to test for differences between disturbed and undisturbed plots.

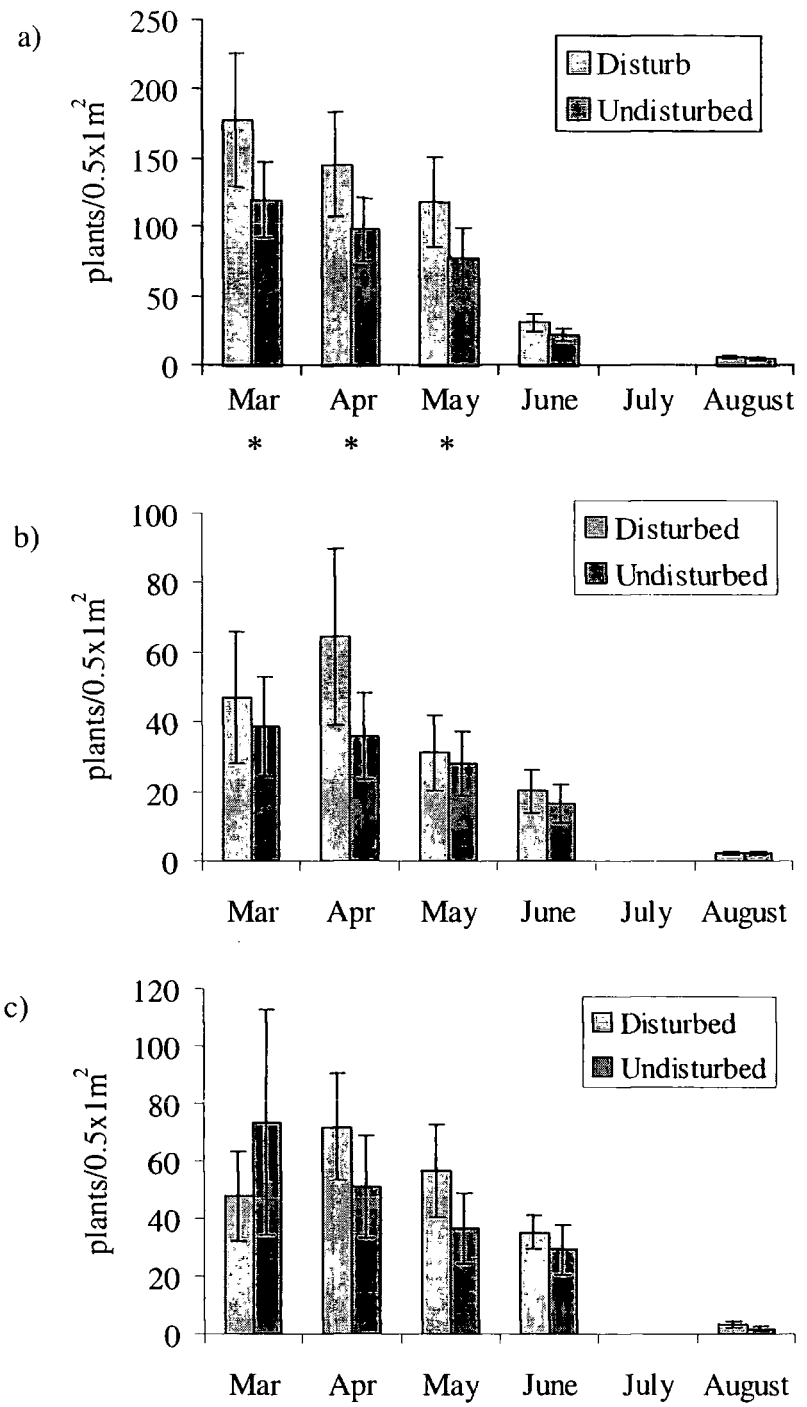

Figure 6.4 Number of *Heracleum* seedlings growing in a) ruderal vegetation, b) woodland and c) grassland in 1998, which had either been previously disturbed or left undisturbed. Pairs of columns marked "*" are significantly different (paired t-test, $P<0.05$), all others are non-significant. Wilcoxon's tests were use in April and March on the woodland and grassland data respectively. No data were collected in July.

The figure indicates that although mean survival of seedlings is almost always greater in disturbed areas, irrespective of habitat, these effects are only significant in ruderal vegetation and then only early in the growth season. By contrast the effects of disturbance on *Impatiens* and *Fallopia* (Figures 6.5 & 6.6) are not significant in any of the habitats nor at any point during the year.

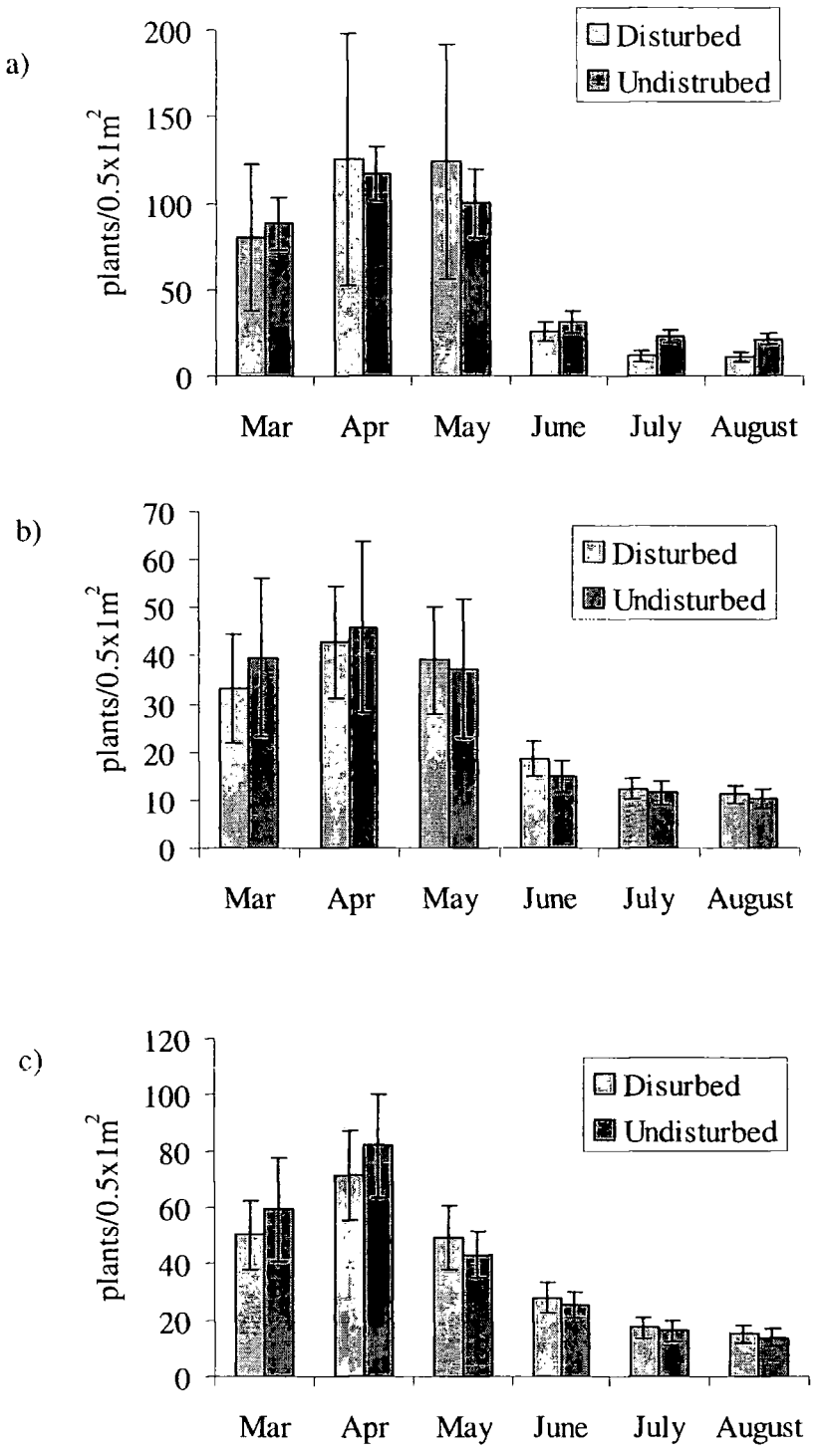

Figure 6.5 Number of *Impatiens* seedlings growing in a) ruderal, b) woodland and c) grassland vegetation in 1998, which had been either disturbed in the previous autumn or left undisturbed. Paired t-tests (Wilcoxon's in Mar-May for ruderal data) are all non-significant.

The emergence and survival of *Fallopia* shoots (Figures 6.6) demonstrates no spring flush of growth followed by competitive exclusion but rather an initial growth of shoots which is slowly augmented by additional shoots appearing during the course of the growth season.

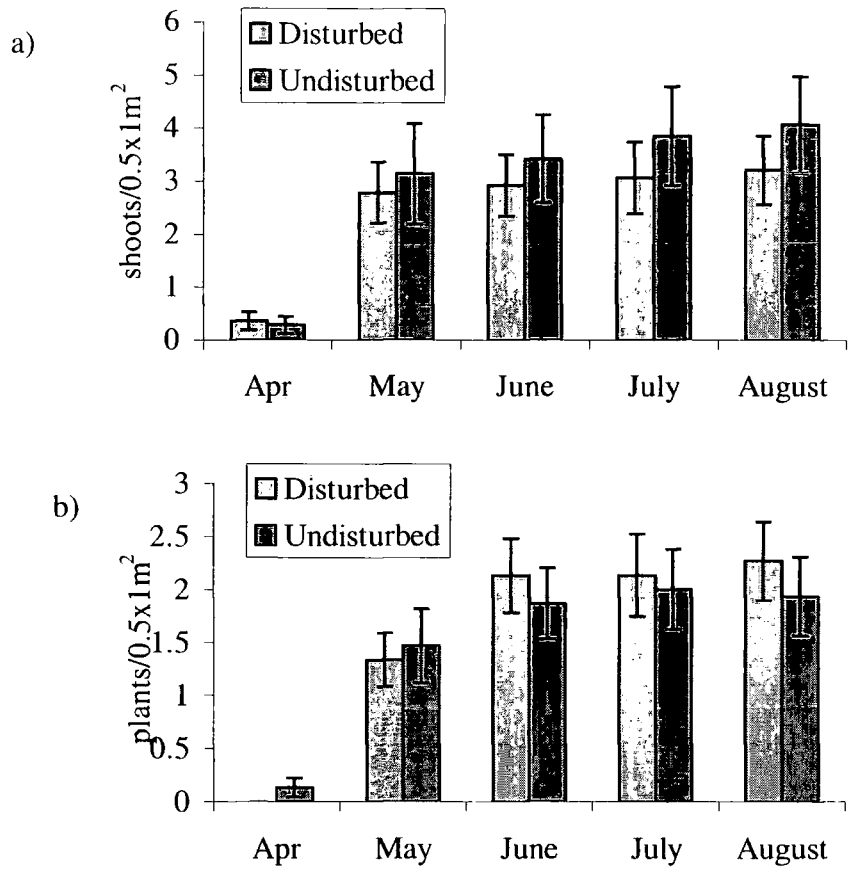

Figure 6.6 Growth and survival of *Fallopia* shoots growing in a) ruderal and b) woodland vegetation in 1998, which had been either disturbed in the previous autumn or left undisturbed. Paired t-tests (woodland in April excluded) are all non-significant.

6.3.3 Growth season mortality

The survivorship graphs of *Impatiens* in undisturbed quadrats in the various habitats (Figures 6.7) highlight differences in the patterns of survivorship in the different years and between habitat types. It can be seen that in general the survival patterns of *Impatiens* in ruderal habitats are similar in the two years of study although there tended to be greater survival to flowering in 1997. This is probably linked to the higher seedling densities in 1998 (see Figure 6.2).

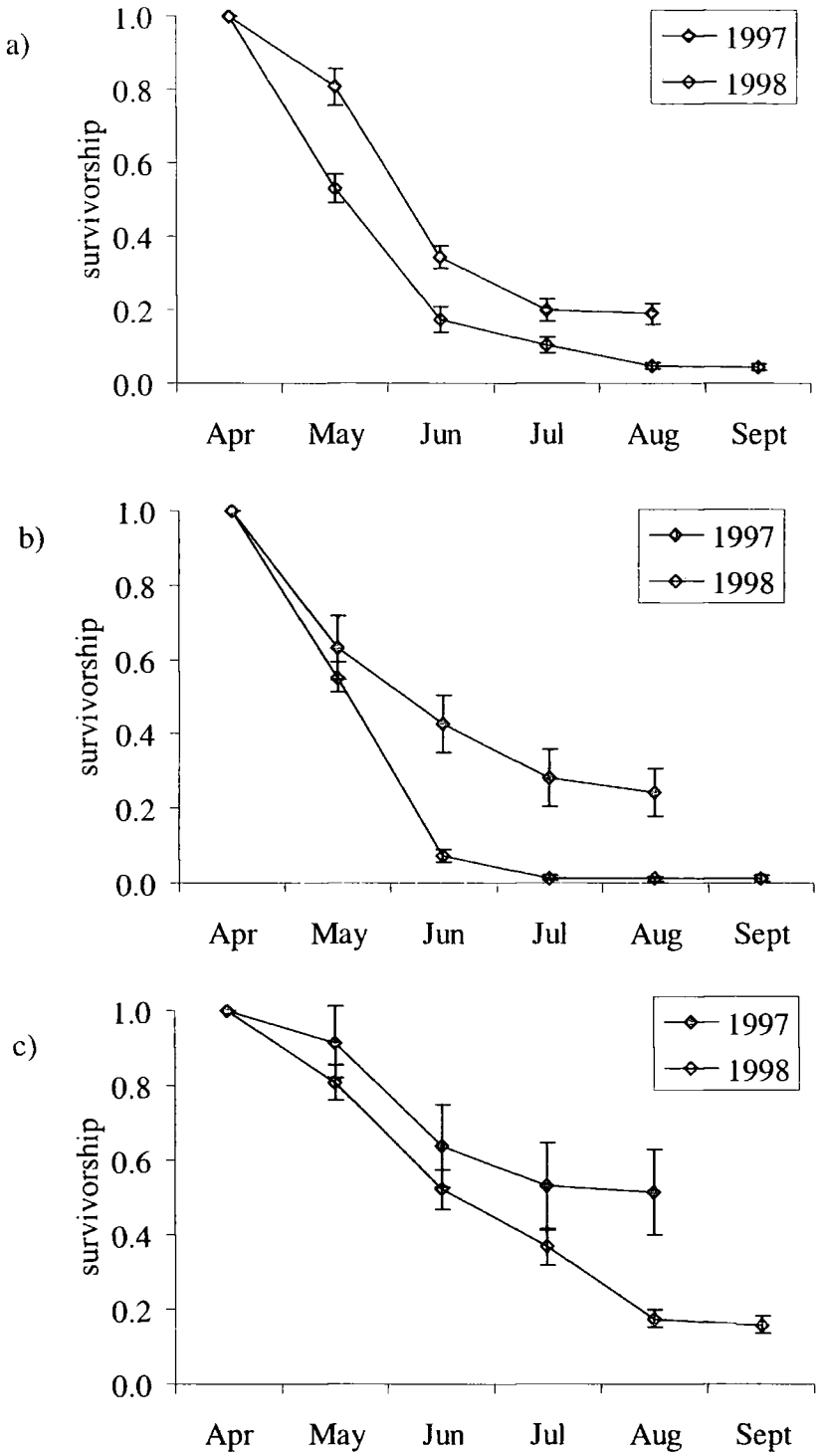

Figure 6.7 Survivorship of *Impatiens* seedlings grown in a) ruderal, b) grassland and c) woodland vegetation in 1997 and 1998.

There were several quadrats in woodland sites in 1997 that showed very little or no mortality over the growing season. These quadrats tended to support only small initial numbers of seedlings, which had high survivorship. The general pattern of survivorship in woodland in both years shows a more continual loss of plants compared to ruderal habitats, which exhibit heavy mortality in the early growing season

The patterns of survivorship of *Impatiens* in grassland vegetation, in contrast to the other two habitats, are markedly different between the two years of study. All of the quadrats monitored in 1998 exhibit rapid initial mortality and subsequent persistence of only very low plant densities. By contrast quadrats in 1997 are typified by very variable surviorship of seedlings, with only a few quadrats showing patterns similar to those observed in 1998.

Life-tables were constructed for plants growing in the different habitats in the two study years. These results are summarised in Figure 6.8, and indicate that there are significant differences in survivorship both between habitats and also between years. In both years mean survival is always greatest in woodland and least in grassland vegetation with ruderal habitats intermediate; though in 1998 these effects are much more pronounced. Trends of differing survivorship between years are not consistent between habitats.

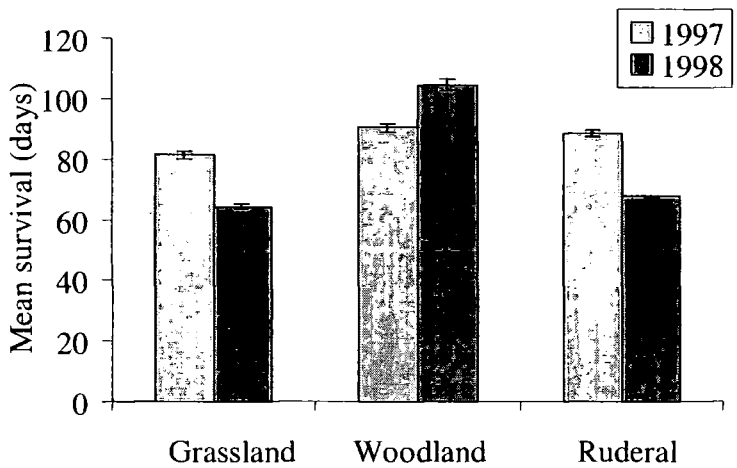

Figure 6.8 Mean survival time (in days) of *Impatiens* individuals in the three habitat types over the growing seasons of 1997 and 1998. Values are derived from Kaplan-Meier survivorship analysis. Log-rank test results of equality of surviorship between i) habitats (adjusted for year) and ii) years (adj. for habitat) both reject the null hypothesis ($S=462.1$, $P<0.001$ and $S=600.7$, $P<0.001$ respectively).

Survivorship charts of *Heracleum* seedlings are shown in Figures 6.9. In several respects these are similar to those shown previously for *Impatiens* in the respective habitats. Like *Impatiens, Heracleum* in ruderal vegetation demonstrates a large mortality early in the growing season before stabilising at lower levels. Similarly in woodland vegetation the two species both show a more constant decline in numbers. Following individual woodland quadrats also highlighted greater variability between quadrats. Survivorship patterns in 1997 generally mirror those shown by the 1998 dataset.

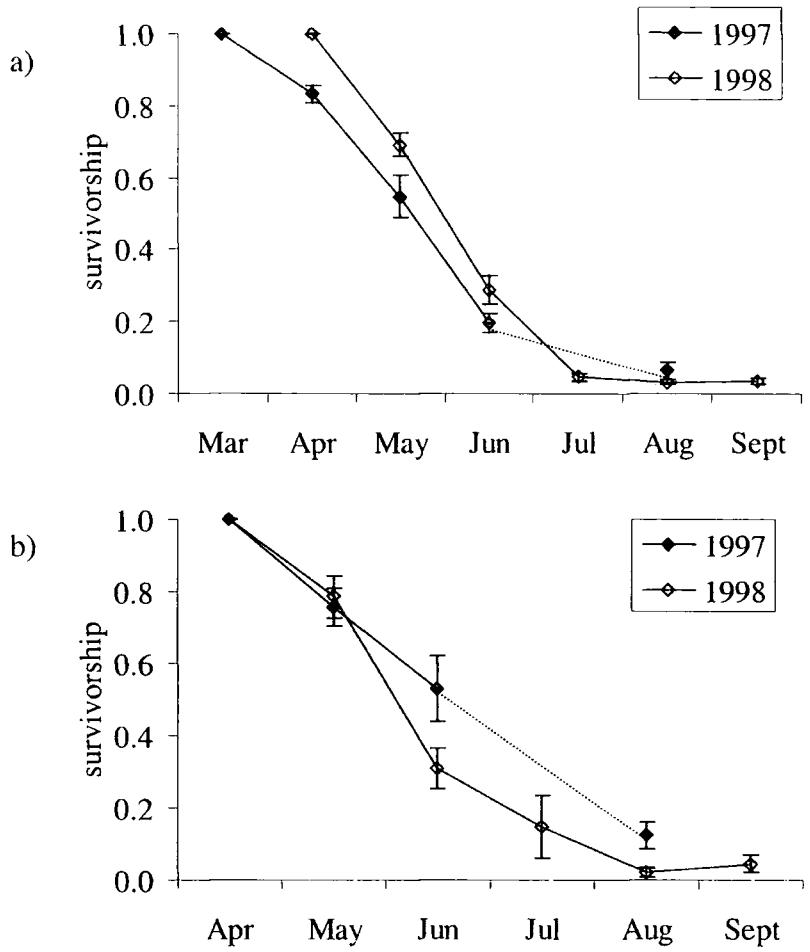

Figure 6.9 Survivorship of *Heracleum* seedlings growing in a) ruderal and b) woodland vegetation during 1997 and 1998. Dashed lines join points with a missing intermediate sampling point.

The survivorship results summarised in Figure 6.10 indicate that although there are no differences in survivorship between years, the effects of habitat do significantly change the survivorship distribution. In both years plants growing in woodland areas have a greater mean survivorship time than those growing in ruderal areas.

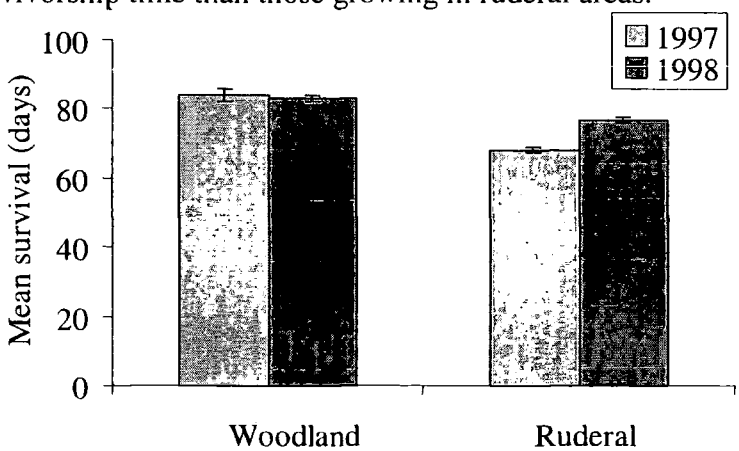

Figure 6.10 Mean survival time (in days) of *Heracleum* individuals in the two habitat types over the growing seasons of 1997 and 1998. Values are derived from Kaplan-Meier survivorship analysis. Log-rank test results of equality of survivorship between habitats (adjusted for year) reject the null hypothesis (S=92.49, P<0.001); whereas results for between years (adj. for habitat) accept the null hypothesis (S=0.131, NS).

6.3.4 Is mortality density dependent?

In order to determine whether the observed low survivorship values of *Impatiens* and *Heracleum* early in the growing season are a result of density-dependent mortality, comparisons were made of plant survivorship in undisturbed quadrats between monthly surveys (Table 6.1). These indicate that there is no habitat in which density dependent mortality is consistently detected for either species in both years of study. *Impatiens* shows evidence of density dependence in all three habitat-types, whereas *Heracleum* displays only a weak relationship of positive density-dependence; the positive slope arising due to a cluster of quadrats which, with low initial densities, resulted in complete mortality (Figure 6.11). Examination of the best of these relationships (*Impatiens* in ruderal vegetation, Figure 6.12) underlines that the relationships are quite weak and that the few quadrats with high plant densities can strongly influence the slope of the regression.

Table 6.1 Summary of significant regression lines derived from plots of "survivorship of plants between a time t and time t+1 against plant densities at time t (where measurements are at monthly intervals)".

Species	Habitat	Year	R^2	Slope	N	F	P
Heracleum	Woods	1998	0.072	+	66	4.970	<0.05
Impatiens	Ruderal	1998	0.1843	-	75	16.497	<0.001
Impatiens	Grassland	1997	0.0877	-	55	5.095	<0.05
Impatiens	Woods	1997	0.1092	-	59	8.301	<0.01

Plots are separated into individual years for each species and habitat type

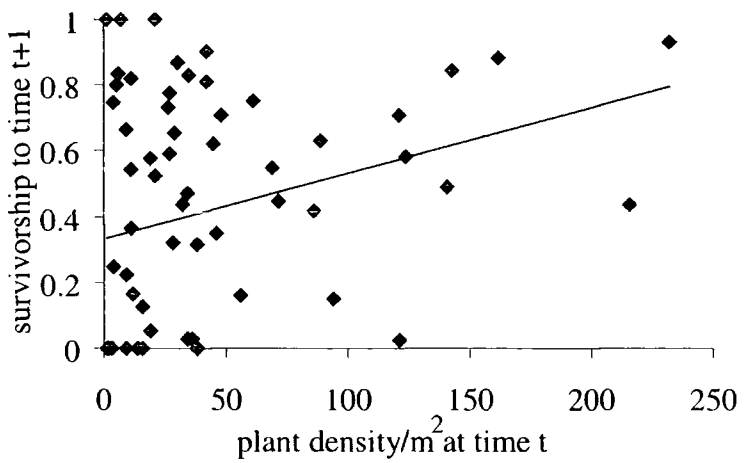

Figure 6.11 Survivorship of *Heracleum* seedlings in woodland (1998) at time t+1 against density at time t, with a four-week recording lag. Equation: $Y=0.0019x + 0.3477$; $R^2=0.072$; N=66, F= 4.97, P<0.05.

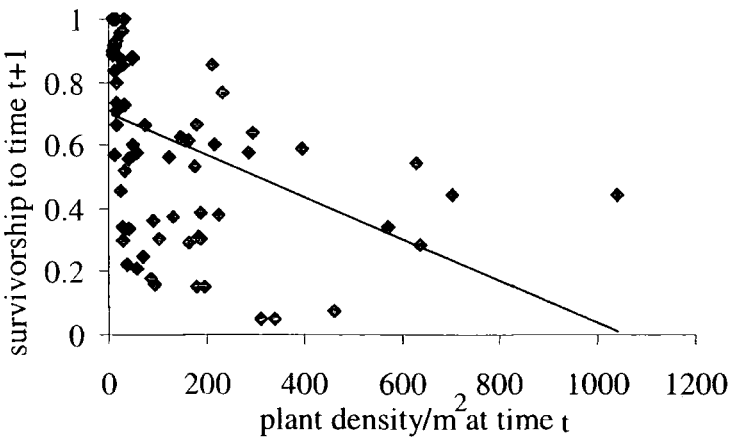

Figure 6.12 Survivorship of *Impatiens* seedlings in ruderal habitat (1998) at time t+1 against density at time t, with a four-week recording lag. Equation Y= -0.0007x + 0.6992; R^2=0.1843; N=75, F= 16.497, P<0.001.

6.3.5 The effects of density-dependent mortality over the growing season

One explanation for the poor fits noted previously could be that similar densities of plants at different times of the year have very different effects. For example 30 *Impatiens* seedlings are unlikely to lead to mortality of other seedlings in a quadrat, whereas 30 plants in mid-summer would cast significant shade and could effectively out-compete underlying plants. As a means of improving the fit of the density-dependent regressions, the quadrat data for each of the species in each habitat were divided into monthly blocks and these points plotted (Table 6.2).

Table 6.2 Summary of significant regression lines derived from plots of "survivorship of plants between a time t and time t+1 against plant densities at time t (where measurements are at monthly intervals)."

Species	dist/ undist	Habitat	Month/year	R-square	slope	N	F	Sig
Heracleum	dist	ruderal	May-97	0.534	-	10	9.38	<0.05
Heracleum	undist	grassland	May-97	0.828	-	5	14.477	<0.05
Impatiens	dist	grassland	Jul-97	0.546	-	13	13.253	<0.01
Impatiens	dist	woods	May-97	0.449	-	12	8.15	<0.01
Impatiens	undist	woods	May-97	0.487	-	15	12.355	<0.01
Impatiens	dist	woods	Jul-97	0.731	-	14	32.574	<0.001
Impatiens	undist	ruderal	May-97	0.299	-	16	5.981	<0.05
Impatiens	undist	ruderal	May-98	0.257	-	15	4.503	<0.05
Impatiens	undist	ruderal	Jun-98	0.591	+	8	8.669	<0.05
Impatiens	dist	ruderal	Jul-97	0.607	-	8	9.273	<0.05
Impatiens	undist	ruderal	Jul-97	0.304	-	15	5.665	<0.05
Impatiens	undist	ruderal	Jul-98	0.638	-	15	22.865	<0.001

Plots are separated into individual months for each species and habitat type. Disturbed and undisturbed quadrats for 1997 results are distinguished.

The significant monthly plots all, bar one exception, demonstrate a relationship of decreasing survivorship with increased density. The one exception suggesting increased

survivorship with density is heavily influenced by two quadrats that had high mortality arising from low initial plant numbers.

Several important features can be suggested from these refined plots. Firstly, it is apparent that *Impatiens* is much more prone to density-dependent mortality than is *Heracleum* (with only two significant regression lines). It also seems that for *Impatiens* density dependent mortality is more likely to occur in areas of ruderal and woodland vegetation than in grassland areas. The R^2 values also indicate much better fits of regression lines once the data are broken up into monthly blocks (see Figure 6.13 below for an example of the improved fit). Lastly, it can also be seen that density dependent mortality occurs only in two months of the growing season, namely between May and June and between July and August.

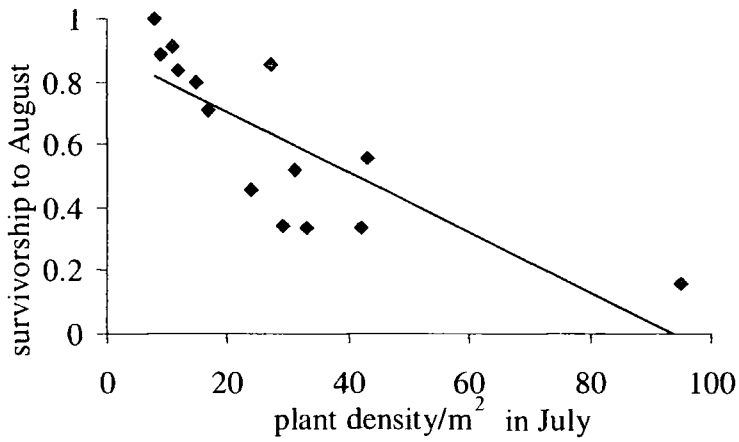

Figure 6.13 An example of survivorship of *Impatiens* seedlings in ruderal vegetation (1998) between July and August against density in July. Equation: $Y = -0.0096x + 0.985$; $R^2=0.6375$, N=15, F=22.865, P<0.001.

6.3.6 Over-winter mortality of *Heracleum*

For those quadrats that remained extant during the study periods it was possible to derive estimates for the over-winter mortality of one-year old *Heracleum* plants and compare this to summer mortality.

The results of summer and winter mortality for *Heracleum* are displayed in Figure 6.14 This suggests that the initial high mortality during the first summer is followed by a period of mortality during the first winter that is also quite severe. These heavy mortality rates during the first twelve months serve to quickly reduce the masses of initial seedlings to low numbers by the following year. In several quadrats the initial flush of seeds were completely eliminated within the first few months.

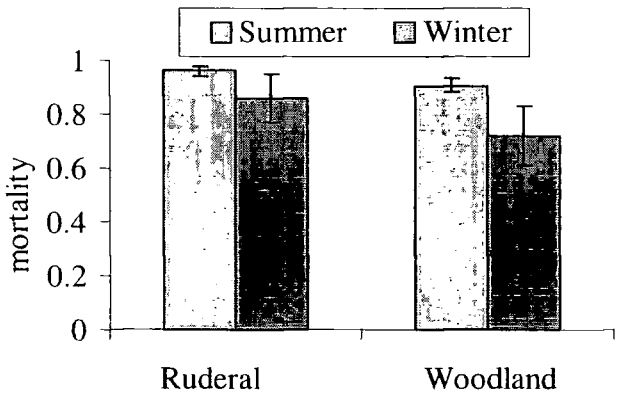

Figure 6.14 Mortality rates for *Heracleum* quadrats remaining extant during the study period. Rates of mortality are shown for summer (the period of a full growing season) and winter (the difference between seedlings present in autumn and 2nd year plants in the following spring). N= 4 & 13 for herb/rud and woodland respectively.

6.3.7 Reproduction

i) *Impatiens*

The results of assessing *Impatiens* seed-pod production in the three habitat types over a three year period indicate quite high variability both between habitats and years (Figure 6.15). In all years, pod production is consistently low in woodland areas. In the other two habitats pod production is lowest in 1996, 1997 producing the biggest crops in ruderal habitats and 1998 producing the best crops in grassland. The high pod production per plant in grassland habitats in 1998 could be due to the low survivorship of plants, leaving more resources for those which survive, though the density of plants at flowering is no different to other years (Figure 6.16). There is a marked consistency in plant densities across all three habitat-types with the lowest densities in grassland areas, higher densities in woodlands and higher still in ruderal vegetation.

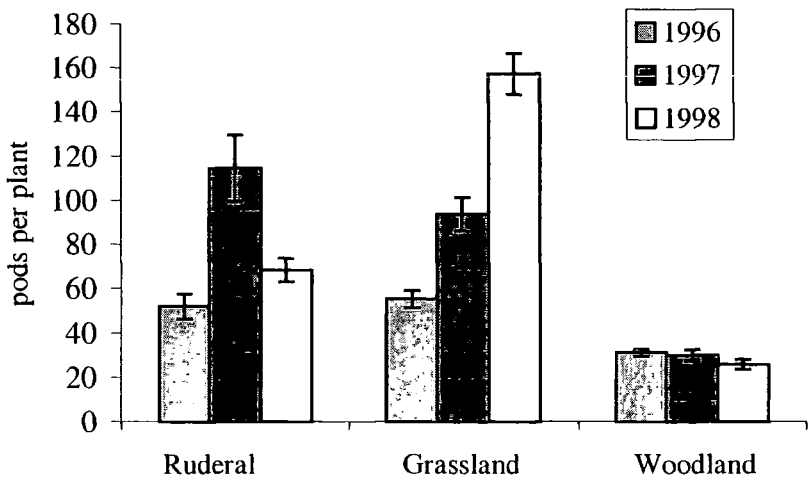

Figure 6.15 Mean seed output of *Impatiens* plants in the different habitats across the three years of the study.

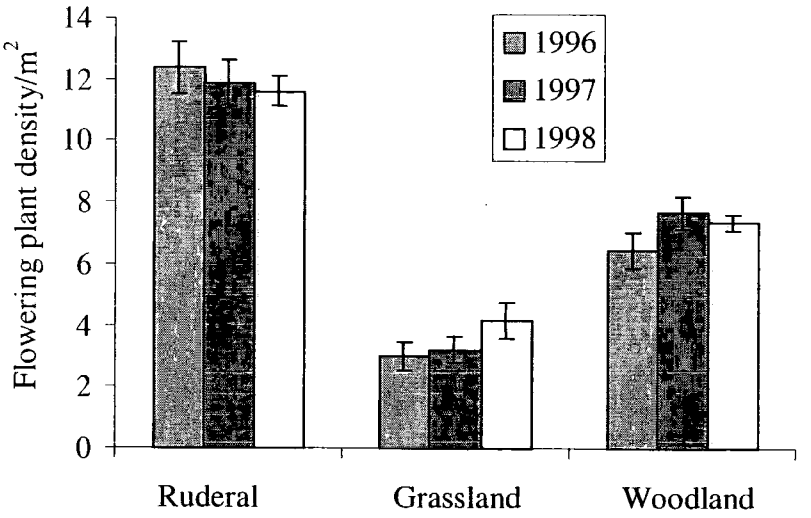

Figure 6.16 Mean density of flowering *Impatiens* plants/m^2 in the three habitat types over the three years of study.

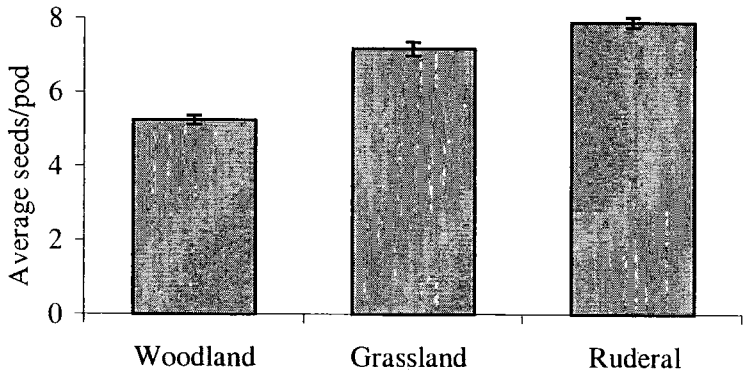

Figure 6.17 Average numbers of seeds in mature pods of *Impatiens* plants growing in the three habitat types. Results are amalgamated from data collected in 1996 and 1997.

The high pod production despite high plant densities in ruderal areas, combined with high seed numbers per pod (Figure 6.17) lead to large seed output per square metre (Figure 6.18). Conversely the low pod production in woodlands combines with low seed production per pod and intermediate plant densities to provide fairly low seed outputs per square metre. In grassland areas the high pod production and intermediate values of seeds/pod are confounded by low plant densities which results in only moderate seed output per square metre.

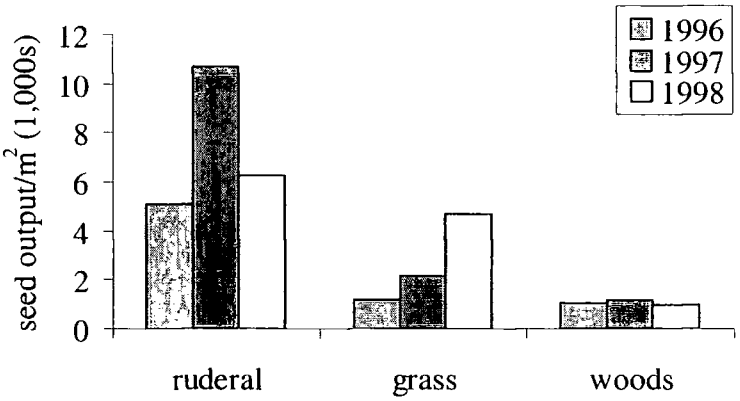

Figure 6.18 *Impatiens* seed estimates per square metre in the three habitats. Values for 1996 to 1998 are derived from mean values of plant density, pod production and seeds per pod

ii) *Heracleum*

The densities of adult *Heracleum* plants are not easily measured at the scale of the study as flowering plants generally occur at densities of 1 plant/m^2 or less. Therefore although the densities from quadrat data between habitats may appear similar, this is likely to be an artefact of positioning quadrats in areas where adult plants occurred. Had the quadrats been of a size of 5m^2 or 10m^2 it is likely that woodland areas would be found to support lower plant densities than ruderal areas.

Plants growing in wooded areas had significantly lower numbers of rays on the terminal umbel than plants growing in ruderal areas (Figure 6.19), a trend which was consistent between years though most marked in 1997. Similarly plants growing in ruderal areas consistently had greater numbers of satellite and axillary umbels than did plants in woodland (Figure 6.20). There was also a significant difference in the number of such secondary umbels produced between years in the two habitats, 1998 being a better year for umbel production.

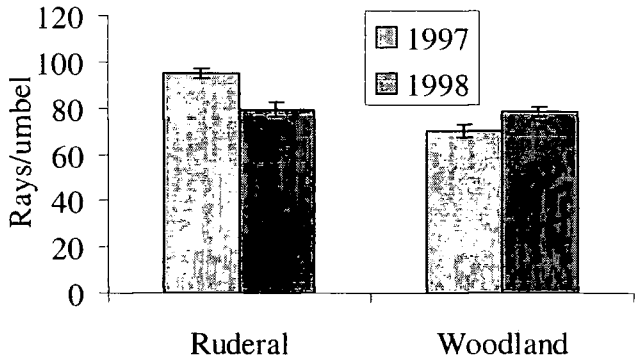

Figure 6.19 Mean number of rays on terminal umbel of plants growing in the two habitat types in 1997 and 1998. ANOVA: year ($F_{1,80}$=1.902) NS; habitat ($F_{1,80}$=22.54) P<0.001; year x habitat ($F_{1,80}$= 19.191) P<0.001.

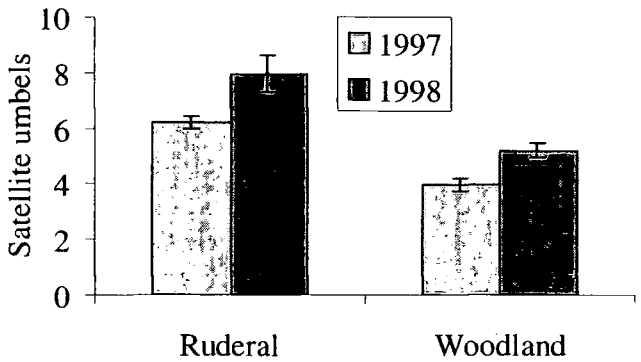

Figure 6.20 Mean number of satellite and axillary umbels on plants growing in the two habitat types in 1997 and 1998. ANOVA: year ($F_{1,80}$ =13.545) P<0.001; habitat ($F_{1,80}$=37.74) P<0.001; year x habitat ($F_{1,80}$= 0.397) NS.

There are significantly more seeds produced on primary umbels in ruderal areas than woodland areas, these effects being most pronounced in 1997 (Figure 6.21). There is no consistent trend in primary umbel seed production between years.

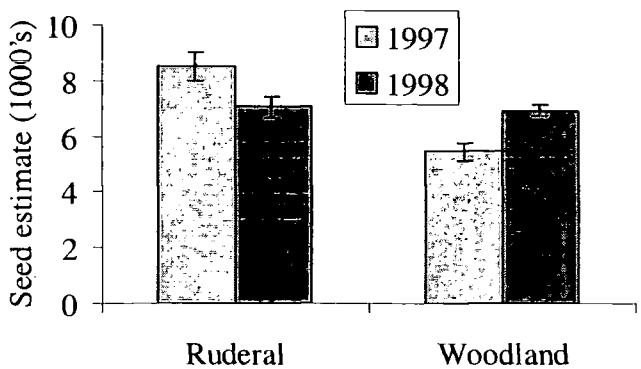

Figure 6.21 Mean number of seeds on primary umbels of plants growing in the two habitat types in 1997 and 1998. ANOVA: year ($F_{1,80}$=0) NS; habitat ($F_{1,80}$=21.36) P<0.001; year x habitat ($F_{1,80}$= 18.22) P<0.001.

The production of seeds by secondary and axillary umbels also shows significant differences between habitats (Figure 6.22). Contrary to the case with the primary umbels, differences in production on these smaller umbels show most difference between habitats in 1998, with a less marked pattern in 1997. This is due, in part, to mass seed failure on axillary and satellite umbels in ruderal areas during 1997, probably as a result of a spell of very dry weather. Hence although the number of such umbels was not too dissimilar between 1997 and 1998, there is a large discrepancy in actual seed outputs. This difference in seed production by these umbels is also significantly different between years in the two habitats.

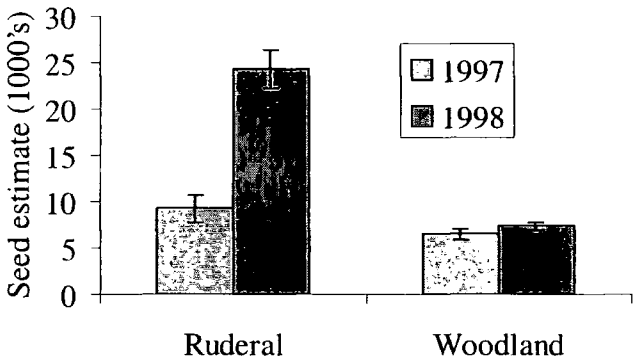

Figure 6.22 Mean totals of seeds on satellite and auxillary umbels of plants growing in the two habitat types in 1997 and 1998. ANOVA: year ($F_{1,80}$=41.472) P<0.001, habitat ($F_{1,80}$=63.873) P<0.001, year x habitat ($F_{1,80}$= 32.831) P<0.001.

A combination of the primary, satellite and axillary estimates for individual plants allow total seed outputs to be calculated and compared across habitats (Figure 6.23). These serve to emphasise what has previously been show for the different umbel classes; namely that there is significantly less seed produced by plants in woodland areas than in ruderal habitat and that there is variation in seed output between years.

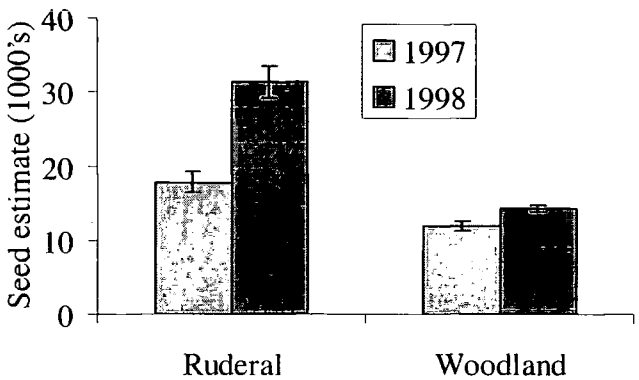

Figure 6.23 Mean totals of seeds produced by plants growing in the two habitat types in 1997 and 1998. ANOVA: year ($F_{1,80}$=36.849) P<0.001; habitat ($F_{1,80}$=76.849) P<0.001; year x habitat ($F_{1,80}$= 18.313) P<0.001.

6.4 Discussion

It is appropriate to relate the various germination, survivorship and fecundity results presented in this chapter to similar figures previously published elsewhere. The mean germination of *Heracleum* propagules in ruderal areas is similar to the value of 400 first leaf stage seedlings m^{-2} reported in Tiley *et al.* (1996), though the values presented here are pre-first-leaf. The seedbank populations of up to 2664/m^2 recorded by Ochsmann (1992) are comparable to the maximum seedlings recorded in ruderal areas during this study, perhaps suggestive of high germination rates. Similarly the densities of seedlings under mid-summer canopies in ruderal sites (e.g. values for June in Figure 6.4-6) are very similar to those of 33/m^2 given by Tiley *et al.* (1996). The latter literature also suggests values of 6-10 established 1-year-old plants/m^2 and mature flowering specimens at densities no more than c. 0.5-1/m^2, which parallels these results very well in both habitat types. The slight caveat mentioned previously, that of siting quadrats next to mature plants, may well give misleadingly high densities of flowering plants in woodlands. This was certainly the impression from visiting sites in the field, lower densities perhaps resulting from *Heracleum* avoiding areas of heavy shading. The densities of flowering plants recorded by Gibson *et al.* (1995) of 4-7/m^2 were not even approached in any of the study sites.

The major differences in the data presented here, compared to elsewhere, are the low germination values observed in woodland vegetation. Later in the season however the woodland plant densities approach those of ruderal areas as a result of increased survivorship.

It was not possible to follow *Heracleum* populations right through from seedlings to flowering (since this would take three or more years), in order to determine mortality rates at all stages of growth. However it was apparent from the collected data that much of the mortality takes place during the period of the first growing season. Following this, a moderate mortality rate over the first winter thins the plants out to such low levels that the mortality in subsequent years must be much reduced to result in the final stand densities of approximately 0.5-1 flowering plant m^{-2}. An additional complication with calculating survivorship curves and mortality figures for *Heracleum* is its ability to die back above ground during a growing season but then re-grow from its rhizome again in subsequent years. Additionally a noted ability to persist for several years bearing just a single leaf (Tiley *et al.* 1996) could lead to confusion between younger cohorts. In reality, if a robust life-table were to be constructed for this species, it would be

necessary to combine permanent quadrats followed over a number of years with individual marking of plants in these quadrats.

Survivorship curves for *Heracleum* indicate a fairly constant drop in numbers in ruderal areas until July, when numbers level off. By contrast, in woodland a whole range of decline patterns were found, though again numbers generally levelled off by July. Mean survival times of seedlings were higher for woodland sites (c. 80 days) than ruderal areas (c. 70 days), with no significant change in survival between years. These survival times may reflect the lower seed densities in woodland, potentially reducing intra-specific competition. The larger variability in survivorship curves in woodland areas may be due to differing inter-specific competition effects. The results do not demonstrate a strong case for density dependent mortality operating on *Heracleum* plants, though there is a suggestion that this may be important in May, a time when the over-wintering plants and seedlings alike are growing rapidly. Tiley *et al.* (1996) suggest that densities of plants in continuous stands appear to be determined by competition for available light in the early part of the growing season, a fact which this work substantiates There is however no suggestion of density-dependent effects operating in woodland sites.

Mortality rates are very high in both habitats during the first summer (90-95%), remaining relatively high over the first winter (70-85%). There seems to be lower over-winter mortality in woodland sites, though the species cited frost tolerance at all stages of growth (Tiley *et al.* 1996) rule this out as being an explanatory mechanism. More likely is the fact that woodland areas are less likely to suffer the inundation events of the lower banks that support ruderal vegetation. These flood events could cause propagule wash-out or burial as well as inundation-related mortality.

The fecundity of the adult *Heracleum* plants determined from this work can also be contrasted with other published results. The numbers of rays on the terminal umbels of plants along the Wear is lower than previous estimates of 120 (Tiley *et al.* 1996). The ray counts of plants growing in ruderal areas most closely approached this figure, those of woodland plants being lower still. Similarly the number of satellite and axillary umbels of plants examined from the Wear in both habitat types (range 4-8) were lower than values of c.20 given in Tiley *et al.* (1996). However, total seed estimates for plants on the River Wear compare well to estimates from several other sources. Values of between 8,050 and 51,200 in ruderal habitats and between 4,000 and 16,600 in woodland areas fall within the range of other workers (see Table 6.3 below). No

sampled plants along the Wear had seed counts approaching the highest values of Tiley *et al.* (1996) and Caffrey. This may reflect sub-ideal climatic or environmental conditions in the study area, or simply the random sampling approach missing the largest specimens.

Table 6.3 Seed estimates for *Heracleum* plants from the literature

Author	Minimum estimate	Maximum estimate	Notes
Pysek (1991)	50,000	50,000	
Neiland (1986)	1,500	18,000	
Brondegaard (1990)	27,000	27,000	
Ochsmann (1992)	1,000	29,000	
Sommier & Levier (1895)	<20,000	<20,000	
Caffrey (unpub.)	1,500	108,000	Tiley *et al.* (1996)
Tiley et al. (1996)	61,762	81,519	Large plants

Published work on the germination of *Impatiens* does not tend to make any distinction between germination in differing habitat types, most workers tending to use populations in ruderal areas. There are few readily available values for observed seedling densities from other sites. Prach (1994) recorded initial densities ranging from 30-70 plants/m^2, low values compared to mean densities in ruderal vegetation along the Wear but more comparable to grassland and woodland areas. Beerling's (1990a) records of densities between 70 and 284 plants/m^2 and the values of up to 350/m^2 by Perrins *et al.* (1990) are slightly higher but these still do not compare to the mean value of c. 500 plants/m^2 in ruderal vegetation in 1998.

Marking of seedlings of *Impatiens* supported the findings of Beerling & Perrins (1993); that virtually all seedlings appear over a period of *circa* 4 weeks. Maximum *Impatiens* seedlings numbers in all habitats occurs in April and a decline to more stable numbers by about June. This concurs with the findings of Prach (1994), who found that plants that survived to mid-May stood a good chance of surviving the rest of the year. The plots of survivorship curves however show that, although initial rapid mortality is the case in ruderal quadrats, in grassland and woodland vegetation the situation is not so straightforward. Quadrats in the latter two habitats show much more variability in their survivorship. This is likely to be a result of the lower initial densities in these habitats, indeed Perrins *et al.* (1993) showed much reduced mortality at lower initial densities. Perrins *et al.* (1993) and Beerling (1990a) indicated that density dependent mortality occurred in the periods from March-July and April to August respectively. Although this study also proves that density dependent mortality does occur in all three habitat-

types, it also shows that this is not always the case and that results from an entire growing season only produce poor fits to linear density-dependent mortality. By examining the survivorship on a month by month basis however the Wear quadrats indicate that there are two main periods of density dependent mortality, between i) May and June and ii) July and August. The most surprising fact from these analyses is that at the start of the growing season when seedlings are most abundant density-dependent mortality is not important. It could be hypothesised that very early in the growing season seedlings are too small to interfere with other seedlings unless growing millimetres apart. By May however the plants are starting to grow true leaves and expand rapidly leading to the observed density-dependent mortality. Similarly the effects in July could be attributed to the intense competition for resources resulting from flowering and side shoots production. The fit of data points to these inter-month mortality slopes are much better than for slopes constructed from the entire datasets. Final plant densities at the study sites are similar to those seen elsewhere (Prach 1994, {$<10/m^2$}, Beerling 1990a, {$50\text{-}70/m^2$}, Perrins *et al.* 1993, {$30\text{-}40/m^2$}, Prowse 1998, $<10/m^2$).

It is interesting to note that although plants growing in woodland and grassland vegetation occur at lower densities than in ruderal areas, it is only in woodland sites that mean survival time is greater than at ruderal sites. Plants in grassland vegetation have lower survival times than those growing in dense ruderal stands. This is perhaps a reflection of greater inter-specific competition or less suitable environmental conditions in grassland communities. Overall summer mortality however is similar in both woodland and grassland vegetation i.e. less than in ruderal areas. This indicates that there is more early-season mortality in grassland than woodland but, by the end of the season, proportional losses are similar.

A comparison of seed production figures, for both *Impatiens* and *Heracleum,* with mortality estimates suggests that, for both species values exceed those necessary for replacement recruitment. This suggests that the finite rate of increase will be greater than one and the invasion criteria will therefore be exceeded, confirming what is already known of the invasive capabilities of these species along waterways.

Seed production of *Impatiens* plants has been shown to be linked to plant density (Beerling 1990a), indeed Beerling produced an equation to link the two factors:

$$\text{Seed Production} = 1886.2 * \text{Plant density}^{(-0.367)}$$

Beerling (1990a) gave pod production values of 140 per plant at 20 plants/m^2 & 95 at 36 plants/m^2, whilst similar work in Germany suggested 250 and 390 pods/plant at densities of 20 plants/m^2 (Koenies and Glavac 1979; quoted in Beerling and Perrins 1993). Although the pod numbers are not directly linked to plant densities in the results presented here, mean pod production can be related to mean plant densities and the figures compared to those above. Despite the plant densities in all three habitat-types being substantially lower than the published densities of 20-40 plants/m^2 the pod numbers are also generally low. Only in ruderal areas in 1997 and grassland vegetation in 1998 do the pod production figures compare to those above; pod production for woodland plants being much less than the published values. Although individual pod counts of plants in the ruderal and grassland areas occasionally approach 200 pods/plant (with a maximum count of 336 pods/plant), on average plants never demonstrated the fecundity of the German plants.

Seed numbers per pod vary between habitat types in the study areas. The value of 5.2 seeds/pod observed in woodland sites is similar to the values of 5.3 and 5.7 found by Beerling (1990). The values for plants in grassland and ruderal vegetation (7.2 and 7.9 seeds/pod respectively) are substantially greater than the published values, even those of the German plants (6.1 and 6.4 seeds/pod). It may then be the case that the plants growing along the Wear compensate for their apparently low pod production with greater seeds per pod. Final seed estimates calculated using mean values of pod production and seed content/pod are mostly within the ranges of 500-800 seeds/plant cited in the biological flora (Beerling and Perrins 1993), though woodland estimates are consistently lower.

The clonal spread of *Fallopia* precludes analysis of factors such as germination, seed production etc. It was thought that some idea of plant demography could be gained from following shoots through a growth season. Unlike the other species however there is no spring flush of growth and subsequent mortality, instead shoot numbers gradually increase over the course of the season thereby ruling out any survivorship analyses. There is a suggestion that plants growing in woodland may appear later than in open areas (Figure 6.6) perhaps reflecting reduced solar radiation.

No quantifiable measures were taken of performance between habitats, however work by Beerling (1990a) has previously shown a reduction of biomass as a result of increased shading.

Maximum shoot numbers suggest that *Fallopia* can produce higher shoot densities (presumably also attaining greater clonal spread) in ruderal areas compared to woodlands. If fragmentation events depend on the volume of plant biomass during the autumn/winter then it follows that ruderal areas will be more important than woodland areas for source populations. This is also combined with the greater tendency for low-bank ruderal areas to experience flood events. Beerling (1990) also noted reduced plant biomass in woodland areas.

The increases in shoot numbers observed between the two study years in continuously monitored quadrats might be considered a sign of rapid expansion of populations. However the fairly constant boundaries of most populations examined in the field over the three-year study period suggest that rapid expansion is probably limited to newly colonised sites; more established colonies only expanding slowly or being limited to fragmentation for further spread. A more rational explanation for the increase is differing climate between the two years. In 1997 most *Fallopia* stands underwent a period of wilting and die-back/browning of shoots, possibly due to dry weather around the time of shoot production. This may well have retarded the production of other shoots during that year. This would also explain the low shoot numbers of 1997 compared to those given in the biological flora etc. (see discussion in previous chapter). Deleterious climatic effects on *Fallopia* have been noted elsewhere and include frost (Beerling *et al.* 1994) and drought (Grime *et al.* 1988; Conolly 1977) intolerance.

An obvious limitation in calculating demographic life-tables for the study species is a lack of seed bank data. Ideally sites would have been sown with known numbers of propagules and their progress followed. Without destructive sampling of the actual quadrat sites it was not possible to know the seed bank size of each quadrat. Without this knowledge it is not possible to say to what extent the initial germination figures are due to seed availability rather than habitat unsuitability. Some approximation of the seed banks of the different habitats could have been gained by sampling in areas adjacent to study quadrats, though this overlooks variability in seed deposition over small distances and in any case, this was precluded by time restrictions.

Generally it might be expected that the woodland and grassland habitats would have smaller seedbanks than ruderal areas due to low plant densities, though flood deposition could alter this. In contrast, though seedbanks of ruderal areas might be expected to be large, they would be more susceptible to the vagaries of flood events. Large between year variations in the size of *Impatiens* stands in the lower zone probably bear testimony

to the effect of scouring flood events. Uncertainty in the longevity of deposited seeds further complicates seed bank estimates.

Other potentially important factors that these experiments have not considered when examining species' demographies include e.g. the effects of soil type, hydrological regime, habitat plant-community composition, potential seed/plant herbivores and local climate. All such factors will have some role to play in unravelling the complex mechanisms operating to determine suitable and unsuitable sites for these species and could prove equally or more informative than the simple habitat-types considered here. What for example is the likely destiny of propagules of these aliens if deposited under a stand of *Petasites hybridus* compared to *Aegopodium podagraria*? There is still much scope for work on the habitat and environmental constraints of these species.

The demographic results presented in this chapter do in fact aid the explanation of variables selected as important from the RCS modelling. The selection of ruderal vegetation as being a good predictor of *Heracleum* and *Impatiens* is validated by the high seed outputs shown by the demographic studies and the preponderance of occurrences in the ruderal lower-zones of the rivers. The inclusion of grassland vegetation as a positive variable for *Impatiens* is substantiated by relatively good seed output values for *Impatiens.* Similarly for *Heracleum,* the fact that this habitat had to be dropped from the demographic studies due to an inability to locate enough sites is testimony to its inclusion as a negative variable. The inclusion of unnaturally disturbed areas as a negative variable in the models for *Impatiens* and *Heracleum* is supported a lack of benefits arising from disturbance, although negative effects were not actually detected. It may be that disturbance during the growing season is the important factor in limiting these species rather than post-season disturbance, which in many respects is simply mimicking flood conditions.

The continual selection of woodland as a variable associated with the species is perplexing given the poor productivity of all the species in such areas. It is likely that woodland fringes are crucial in harbouring such species and away from such areas their spread should be much less prolific.

One of the most important findings arising from the work presented in this chapter is that much currently published data for these species deal solely with occurrence and performance in optimal habitat areas. The variation in performance between habitats noted here has until now remained largely unreported, though it is likely to have profound influences on modelling the spread of these species.

Chapter Seven

Species Traits along a Climatic Gradient

7.1 Aim

Statistical models suggest that a major constraint on the distribution of invasive alien species is climate (Beerling *et al.* 1995; Beerling 1993; Chapter 3), which may explain their relative absence in habitats of the upper catchment. The specific aim of this component of the work was to examine how elevation, acting as a proxy for climate, influenced the performance of the three species. A series of experiments was set up at various altitudes along one river catchment. The River Wear catchment was chosen due to its altitudinal range and also to complement the surveying work that was conducted on the catchment. Several interpolated climatic variables were available (Barrow *et al.* 1993), as thirty-year mean values on a 10km grid, which could be utilised to provide background information on the climatic variation between sites. Additionally, current meteorological records were available close to two of the sites, which could be used to further interpolate climatic variables related to the exact period of study.

This component of experimental work aimed to substantiate whether the species distributions were currently limited by climate or whether other factors such as dispersal or introduction history were more important. In addition it was hoped to highlight the exact mechanism by which any climate variables operate to restrict the species distributions.

7.2 Methods

7.2.1 Site locations and layout

Five sites were chosen at locations along the Wear (see Figure 7.1). An additional site was also located at Great Dun Fell in Cumbria to encompass a harsh upland climate. The sites were chosen so as to be representative of the range of altitudes, and hence climates in the county (see Table 7.1) and included areas of the catchment in which the species were already present as well as sites from which they were absent.

Table 7.1 Variation in altitude, climate and geomorphology of the chosen study sites (data derived from Graham 1988; Barrow *et al.* 1993).

	Washington	Durham	Wolsingham	Westgate	Rookhope	Great Dun Fell
Altitude (m)	10	30	130	270	330	600
Mean annual pptn (1916-50) /mm	~700	~765	~890	~1145	~1145	~2000
No. of days with temp. over 5^0C (1961-90)	157	151	136	112	109	73
Geomorph. region	Lowlands	Lowlands	Plateau	Plateau	Uplands	Uplands

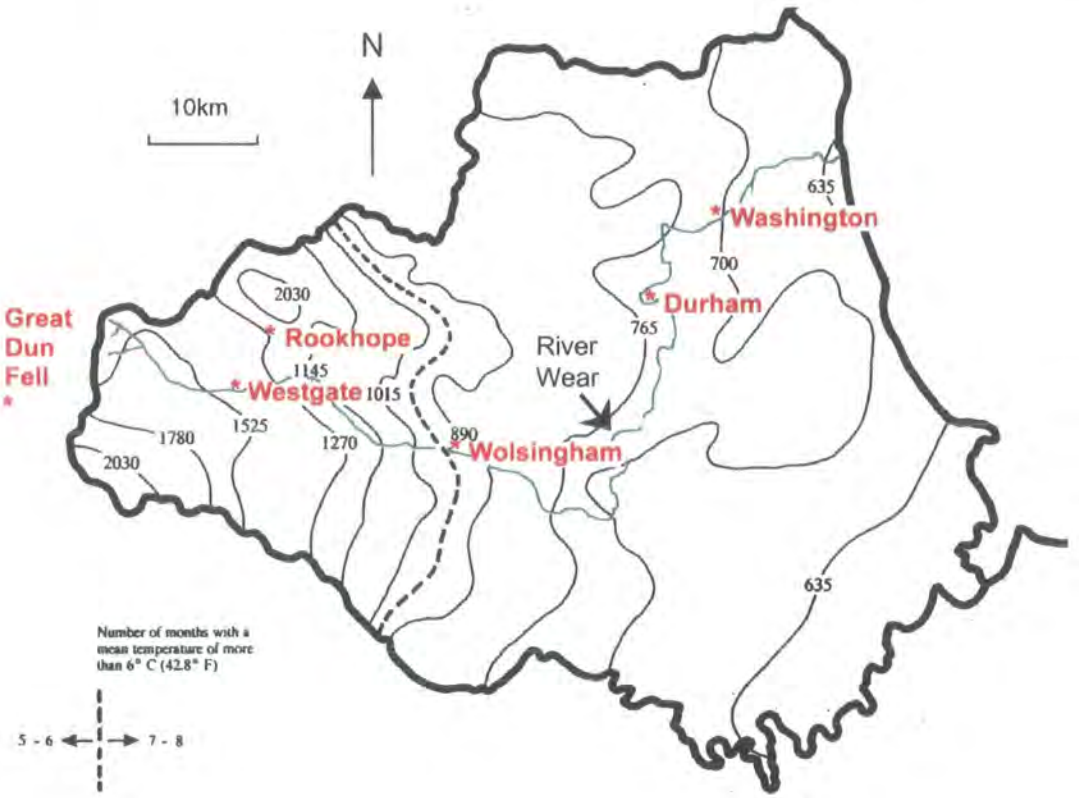

Figure 7.1 Locations of the five experimental sites located along the Wear catchment and the additional upland site on Great Dun Fell. Also shown on the map are the mean annual precipitation values (mm/yr) for Durham and the boundary between the areas with 5-6 months with an mean temperature >6^0C and those with 7-8 months. Adapted from Graham (1988).

7.2.2 Experimental design

At each of the stations several plant demographic traits were monitored. For the seed producing species these included germinability, seedling survival, growth, flowering and seed set. For *Fallopia* traits such as shoot production from rhizomes, growth, flowering and biomass production were monitored.

Because of the laws regarding the deliberate introduction of these species into the wild (Wildlife and Countryside Act 1981) it was necessary to devise an experimental protocol which allowed estimates to be made of plant characteristics without releasing the species into the natural environment. Therefore plants were grown in containers at each site, mimicking the natural environment as much as possible, and propagules were collected before they had been dispersed.

Each individual site consisted of thirty individual pots, ten for each of the three species. The pots used were standard florist's buckets of 25cm diameter and 25cm depth, with drainage holes drilled into the base of each pot. The pots were filled to within 2cm of the rim with standard potting compost and planted with the seeds/rhizomes in a regularly spaced order. For *Impatiens* and *Heracleum*, twenty seeds were sown in each pot at a depth of 1cm whilst for *Fallopia* five rhizome fragments of a standard size (approximately 7cm long and 10mm diameter) were planted at a depth of 3cm in each pot.

Collection of the propagules for the experiment was undertaken in the autumn they were due to be planted. For *Heracleum* and *Impatiens* ripe seeds were collected from adult plants and then stored in paper bags at 5°C until early winter (the beginning of November), at which point they were planted in the pots and left to over-winter at the altitudinal sites. *Fallopia* rhizomes were collected on the day before they were to be planted and cut to the desired length just before planting to ensure rhizomes did not dry out. This also took place in early winter, planted pots being left on site over winter.

Species treatments were distributed amongst the 30 pots at random. A wooden frame, of the same height as the pots, was built to surround the pots at each site. A plastic sheet was used to line the ground within the frame and also extended up the edges of the frame. This acted as a means of water retention and also prevented the escape of rhizomes from drainage holes in the pots. A few drainage holes were put into the plastic sheeting between pots to prevent waterlogging. Around this frame was placed a rabbit proof fence to a height of 75cm, the base being buried or pegged to prevent entry by rabbits. Within the frame the pots were embedded in sand to the height of the soil level in the pots. This ensured conditions within the pots mimicked natural conditions in the soil as much as possible; reducing side-wall heating of the pots and excess evaporation of moisture from within the pots.

In order to allow *Impatiens* plants to flower naturally, flowers were left open to natural pollinators. However to prevent seed dispersal, once seedpods were produced they were enclosed in muslin bags that were then collected after the seeds had been shed. This allowed an estimation of seed production to be made.

Plant performance was assessed every two-weeks. The numbers of germinating seedlings/shoots produced were noted along with a record of their morphological stage (seedling, true-leaf stage, flowering etc.). Additionally a record of plant height was taken, being the length of the straightened stem of a plant including the length of the terminal leaf (or the terminal inflorescence for *Impatiens*). Seed pod and flower production were also recorded for each *Impatiens* plant. *Fallopia* and *Impatiens* plants were monitored over one growing season and then harvested, whereas the *Heracleum* plants from 1997 were left to over-winter and grow for a second season.

At the completion of the first study season (Autumn 1997) all of the *Fallopia* pots were removed and taken into the lab where destructive sampling assessed biomass production during the growing season. Recorded measures included the weight of the initial planted rhizome (weighed in the autumn of harvesting), above ground biomass and excess additional below ground biomass i.e. that excluding the initial rhizome weight. All weights were taken as fresh weights. Biomass recording of *Impatiens* was not undertaken as the plants were in varying stages of wilting/decay on completion of flowering.

After the first growing season the pots were prepared for replanting for the second year of the experiment. Pots containing the original *Heracleum* plants were left untouched apart from topping up the pots with compost and some removal of weeds. To permit a second season of germination experiments with *Heracleum* seeds it was necessary to add another ten pots to each of the sites (after extending the frames). These were filled with new compost and planted out with fresh *Heracleum* seeds, collected from plants growing in the same area from which seeds were gathered the previous year. Similarly the pots that had contained *Impatiens* and *Fallopia* were re-filled with compost and seeds/rhizomes added as in 1996.

The *Heracleum* plants were left to grow for a second year in the hope that some flowering and seed production would be observed. However during the second year of growth these plants increased their biomass to such an extent that they threatened to shade out species in adjacent pots. As it was clear that flowering was not going to occur

(with the exception of one plant) all plants were removed and destructively analysed in late May. The one plant showing signs of flowering was left until seed development occurred before its removal.

During the 1998 growing season *Impatiens* seeds were collected from plants on a regular basis through the summer. This allowed estimates of pod size to be monitored over time along with seed weights, seed being weighted on the day of collection using a sensitive balance.

7.2.3 Climate variables

i) Interpolated climates

Interpolated climate variables were derived for each study site from the 30-year mean values produced by Barrow *et al.* (1993) as part of the Terrestrial Initiative in Global Environmental Research (TIGER) framework. This data-set provides a variety of climate variables (e.g. precipitation, maximum/minimum temperature, growing degree days) for minimum, mean and maximum altitude of all 10km grid squares in the UK. Estimated values for the study sites were obtained by making a linear interpolation of climatic values given for altitudes above and below a site in a 10km grid-square, the interpolation being based on altitude.

Selected examples of derived climate values for the sites at different times of the year are given in Figure 7.2 This shows a trend of decreasing maximum and minimum temperature with increasing altitude, the only exception being the lowland site at Washington where growing season maximum temperatures are slightly lower than the more inland site at Durham. Precipitation is shown to increase with site altitude, whereas wind-speed is usually lowest midway across the transect and higher on the coast and in the uplands. However in mid-summer wind-speed changes to a trend of decreasing speed with increasing altitude.

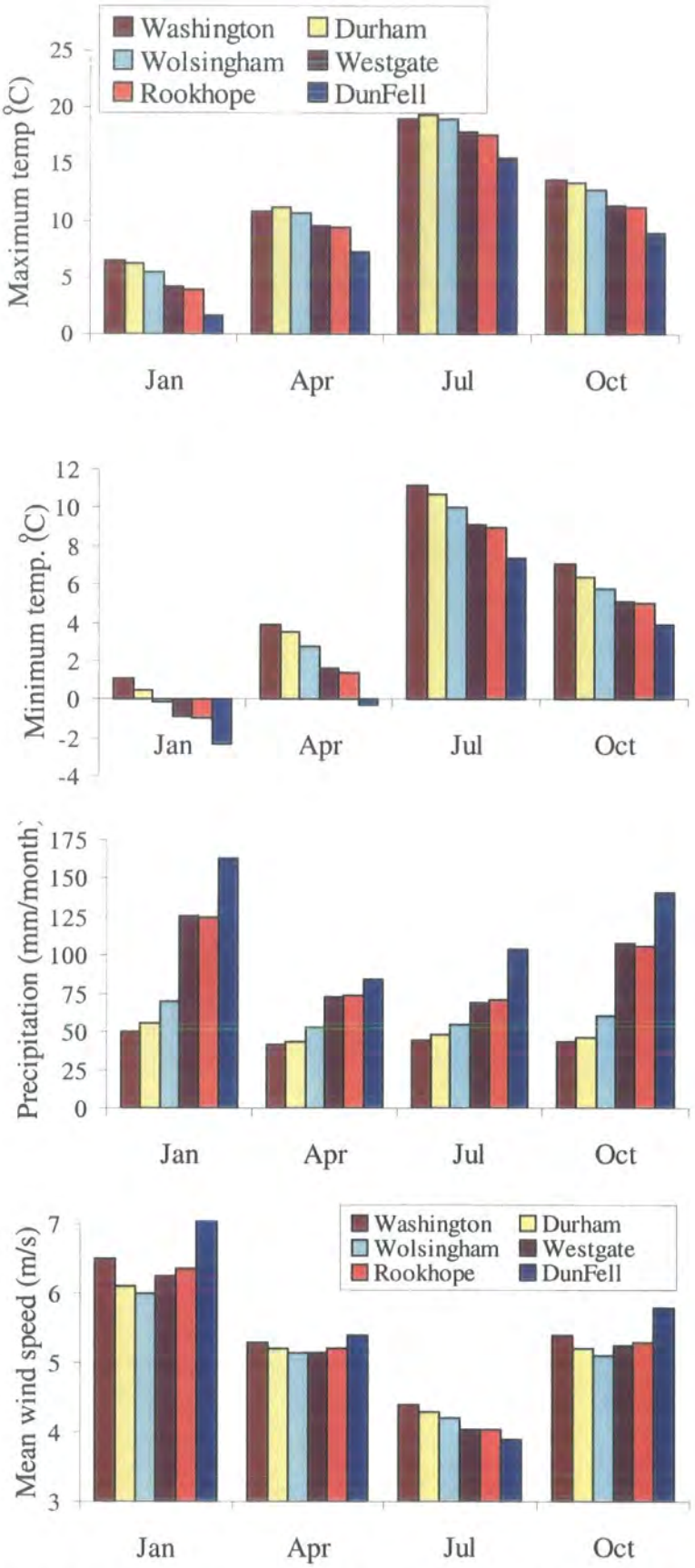

Figure 7.2 Selected climatic data extracted from the TIGER climate dataset for the altitudinal study sites. Values shown are means for each site during the months of January, April, July and September.

ii) Meteorological data

Climate data over the period October 1996-October 1998 were obtained from a standard meteorological station at Durham and an automated station on Great Dun Fell. From these it was possible to derive daily maximum, mean and minimum temperatures along with daily precipitation and wind-speed over the experimental period.

These data were then used in conjunction with the interpolated 30-year climate data (Barrow *et al.* 1993) to simulate daily maximum and minimum temperature for the other sites where temperature observations were not available. This was done using an interpolation method similar to that described by Parker *et al.* (1992) but additionally encompassing the available 1961-90 interpolated climate dataset.

Using the Durham Observatory data, monthly maximum and minimum temperature anomalies were calculated for the entire period. These anomalies were then used in conjunction with the interpolated 1961-90 means to derive a monthly mean maximum and minimum temperature for the sites without meteorological data. At a monthly time step, the daily maximum and minimum temperature at Durham and Great Dun Fell were separately converted to standardised anomalies (z-scores) by subtracting the monthly mean and dividing by the standard deviation of each series. The mean value of the resulting daily z-scores and monthly standard deviations for these two sites are then used to calculate a regional measure of relative daily temperature variability for the remaining sites using an inverse distance-squared weighting. Finally daily maximum and minimum temperatures are modelled for each of the sites by multiplying the daily z-scores by the relevant monthly standard deviation and adding the modelled monthly mean from the TIGER dataset. From these values it is then possible to calculate parameters such as frost-days and growing degree-days for each of the sites over the study period (Figure 7.3).

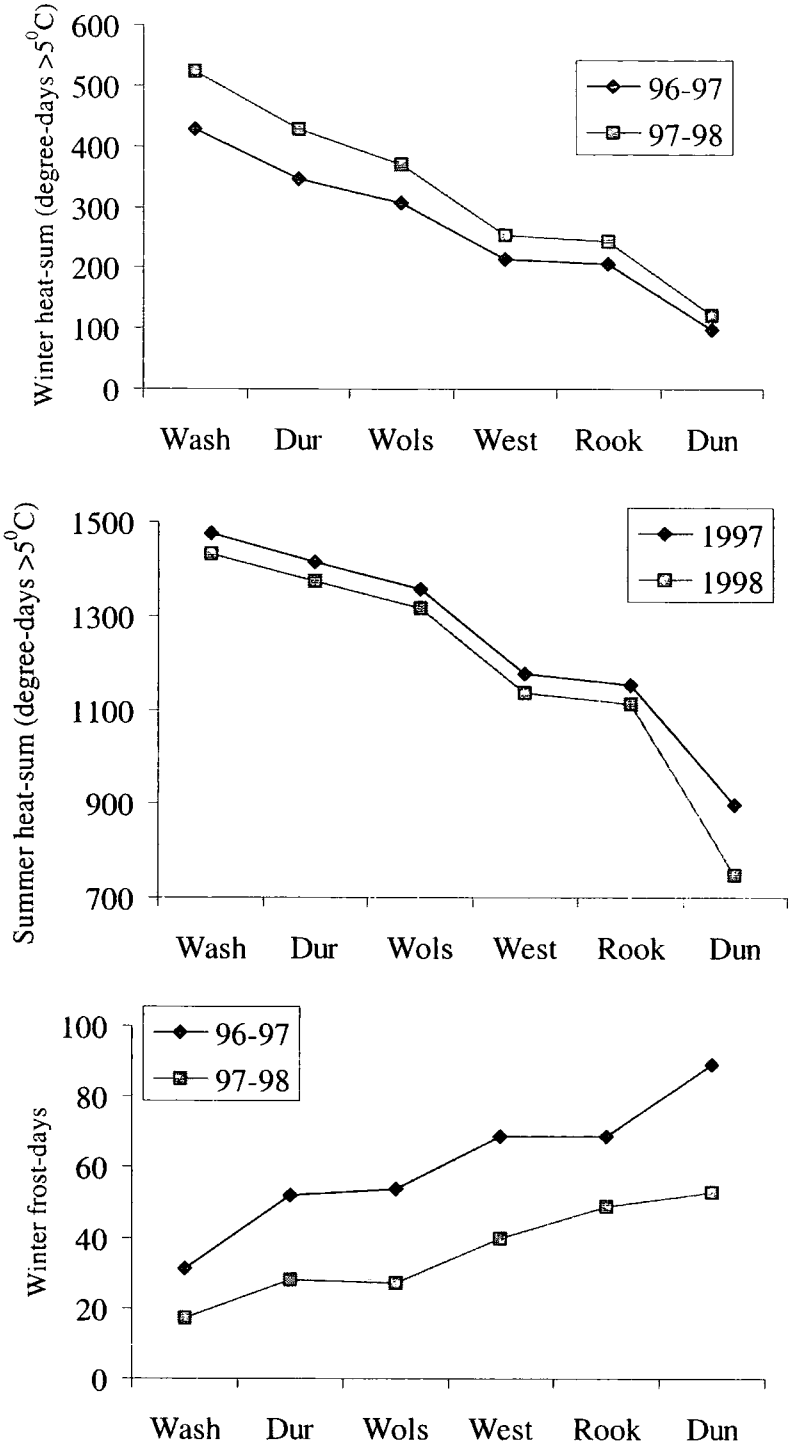

Figure 7.3 Climatic parameters for the transect sites over the study period derived from the interpolated data-sets. Graphs show a) winter heat-sum, b) summer heat-sum and c) winter frost days.

7.3 Results

7.3.1. Germination

a) Timing of germination

Germination of seeds of *Impatiens* and *Heracleum* and growth of shoots for *Fallopia* are shown, in terms of timing, in Figures 7.4. There was a significant difference in the timing of "germination", both between sites in any year (*Impatiens*: $F_{1003,5}=26.30$, $P<0.001$; *Heracleum*: $F_{1962,5}=160$, $P<0.001$; *Fallopia*: $F_{326,5}=13.65$, $P<0.001$) and also between years (*Impatiens*: $F_{1003,1}=563.9$, $P<0.001$; *Heracleum*: $F_{1962,1}=8251$, $P<0.001$; *Fallopia*: $F_{326,1}=123.1$, $P<0.001$) , for all three species.

A significant interaction between the two factors for all three species (*Impatiens*: $F_{1003,5}=27.11$, $P<0.001$; *Heracleum*: $F_{1962,5}=65.94$, $P<0.001$; *Fallopia*: $F_{326,5}=9.79$, $P<0.001$) suggests differing patterns of between-year germination at the sites. With *Impatiens* the trend of later germination date with increasing altitude shown in 1997 disappeared in 1998. In the case of *Heracleum* there was also a general trend of later germination with increased altitude in 1997 which became less pronounced in 1998, only the highest site having a notably later germination date. Conversely the growth of *Fallopia* shoots in 1997 showed no trend with increasing altitude, whereas in 1998 a general trend of later germination with altitude was apparent.

Comparing the three species it can be seen that *Impatiens* seedlings appear consistently earlier than *Heracleum* seedlings which in turn usually appear before *Fallopia* shoots.

A comparison of the germination date of propagules of the species with climatic estimates identified two variables as potentially being important in determining timing of germination. There were significant trends for later germination of all three species with increased incidence of spring frost days (Figure 7.5) and spring heat-sum (Figure 7.6). For all three species the spring heat-sum variable proves a better explanatory variable of germination date than does spring frost incidence. The fit of the regression lines are improved somewhat if mean values for site/year (not shown) are used rather than the data for individual plants, suggesting there is much variability between plants within a site.

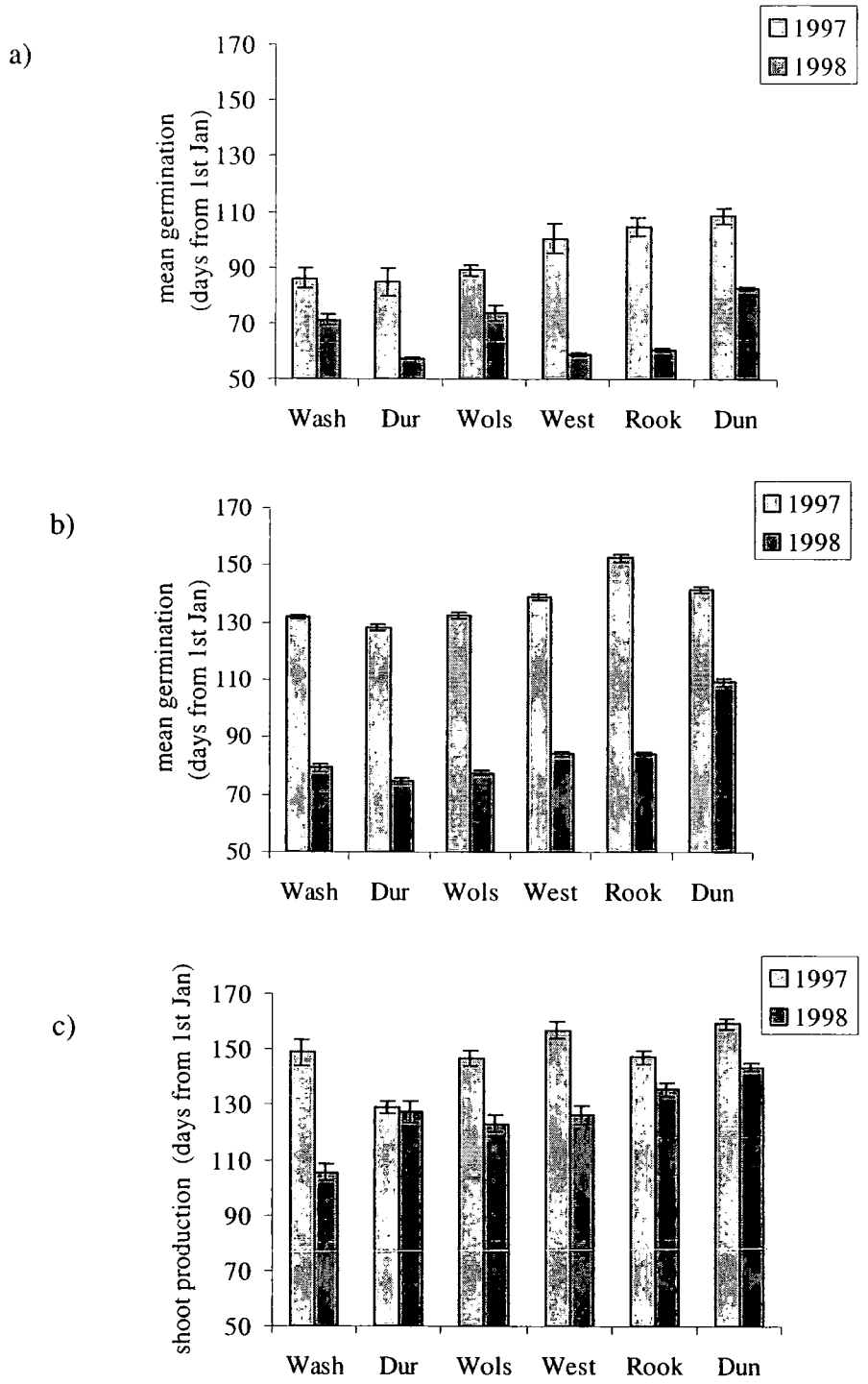

Figure 7.4 Mean germination date (measured in days from January 1st) of a)*Impatiens* seedlings, b) *Heracleum* seedlings and c)*Fallopia* shoots grown in the experimental plots along the altitudinal transect.

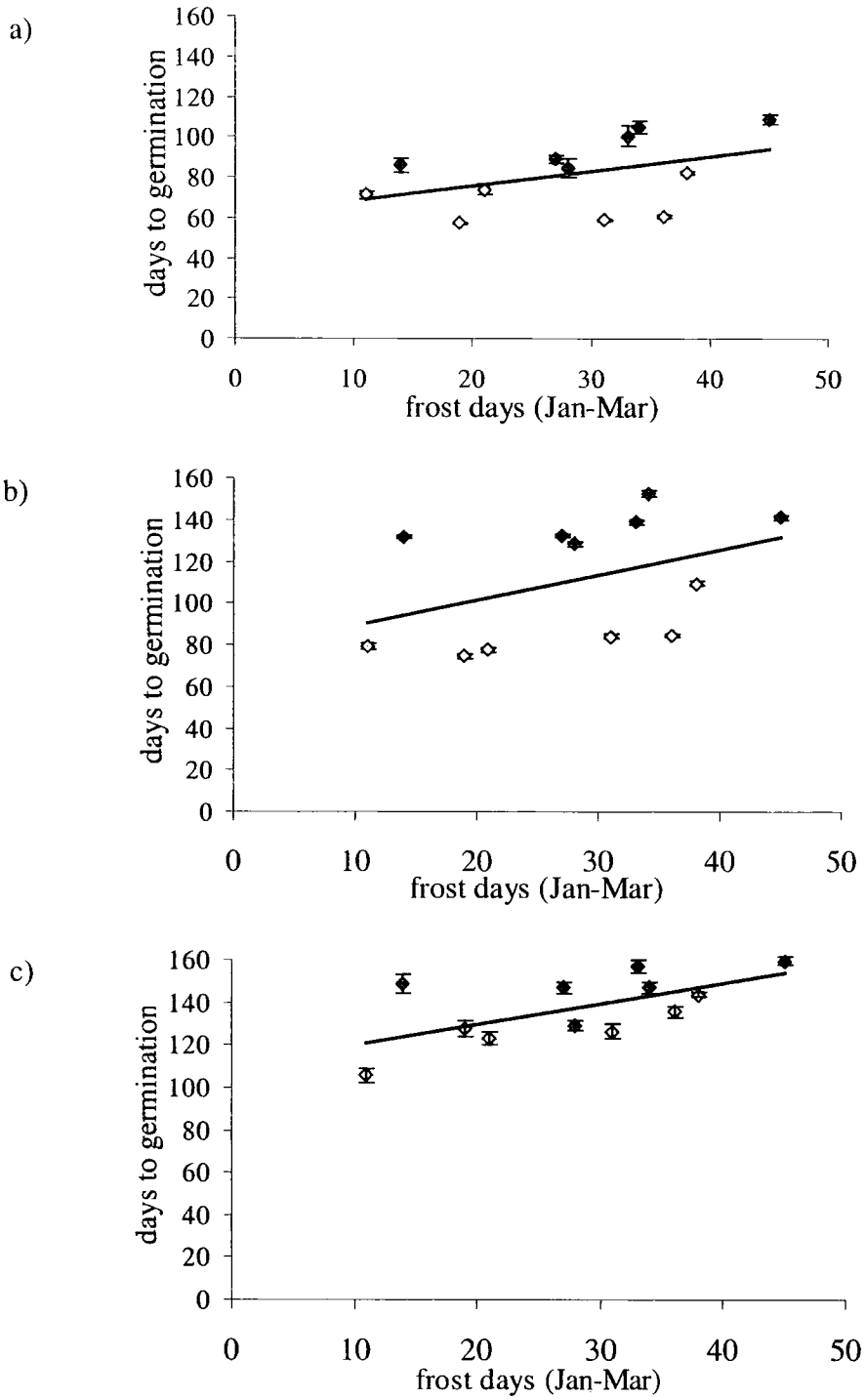

Figure 7.5 Germination time (measured as days from 1st Jan) of seedlings of a) *Impatiens* and b) *Heracleum*, and c) shoots of *Fallopia* in relation to Spring frost days (Jan-Mar) at a site. Data are derived from plants grown in the altitudinal experiment during 1997 (filled) and 1998 (open). Displayed values are means (with S.E.) from each of the study plots in a particular year. Regressions calculated using all data (not means).

Species	Equation	r^2	F value	Signif.
Impatiens	Y=0.397x + 57.69	0.043	$F_{1013,1}$=45.527	P<0.001
Heracleum	Y=1.157x + 79.239	0.134	$F_{1972,1}$=304.50	P<0.001
Fallopia	Y=0.862x + 114.99	0.126	$F_{336,1}$=48.572	P<0.001

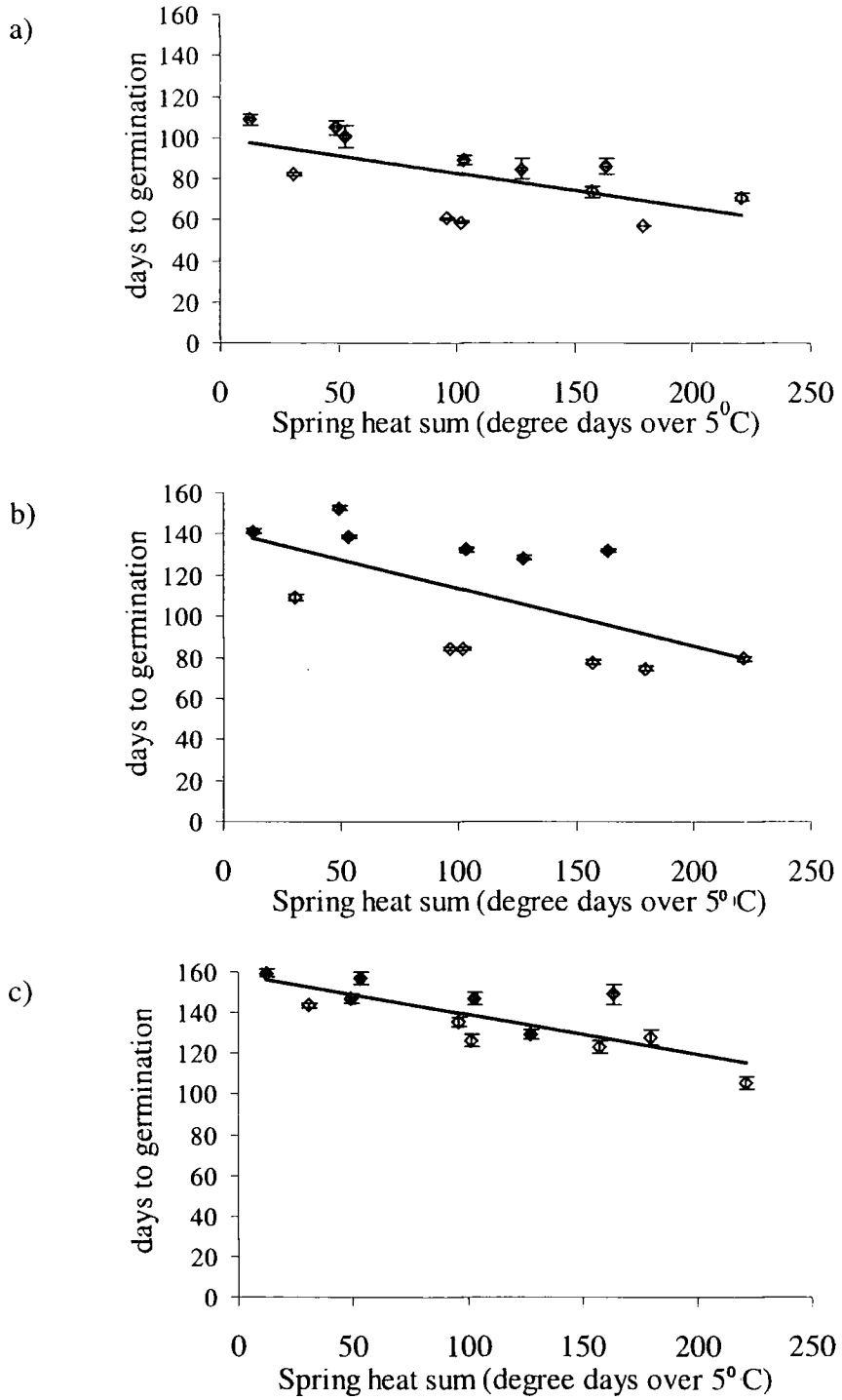

Figure 7.6 Germination time (measured as days from 1st Jan) of seedlings of a) *Impatiens* and b) *Heracleum*, and c) shoots of *Fallopia* in relation to Spring heat sum (Jan-Mar) at a site. Data are derived from plants grown in the altitudinal experiment during 1997 (filled) and 1998 (open). Displayed values are means (with S.E.) from each of the study plots in a particular year. Regressions calculated using all data (not means).

Species	Regression	r^2	F-value	Signif.
Impatiens	Y=-0.112x + 81.87	0.149	$F_{1013,1}$=177.77	P<0.001
Heracleum	Y=-0.277x + 141.64	0.301	$F_{1972,1}$=850.6	P<0.001
Fallopia	Y=-0.179x + 157.91	0.231	$F_{336,1}$=100.79	P<0.001

b) Percentage germination

There were differences in the amount of "germination" of the three species between sites and between years (Figure 7.7).

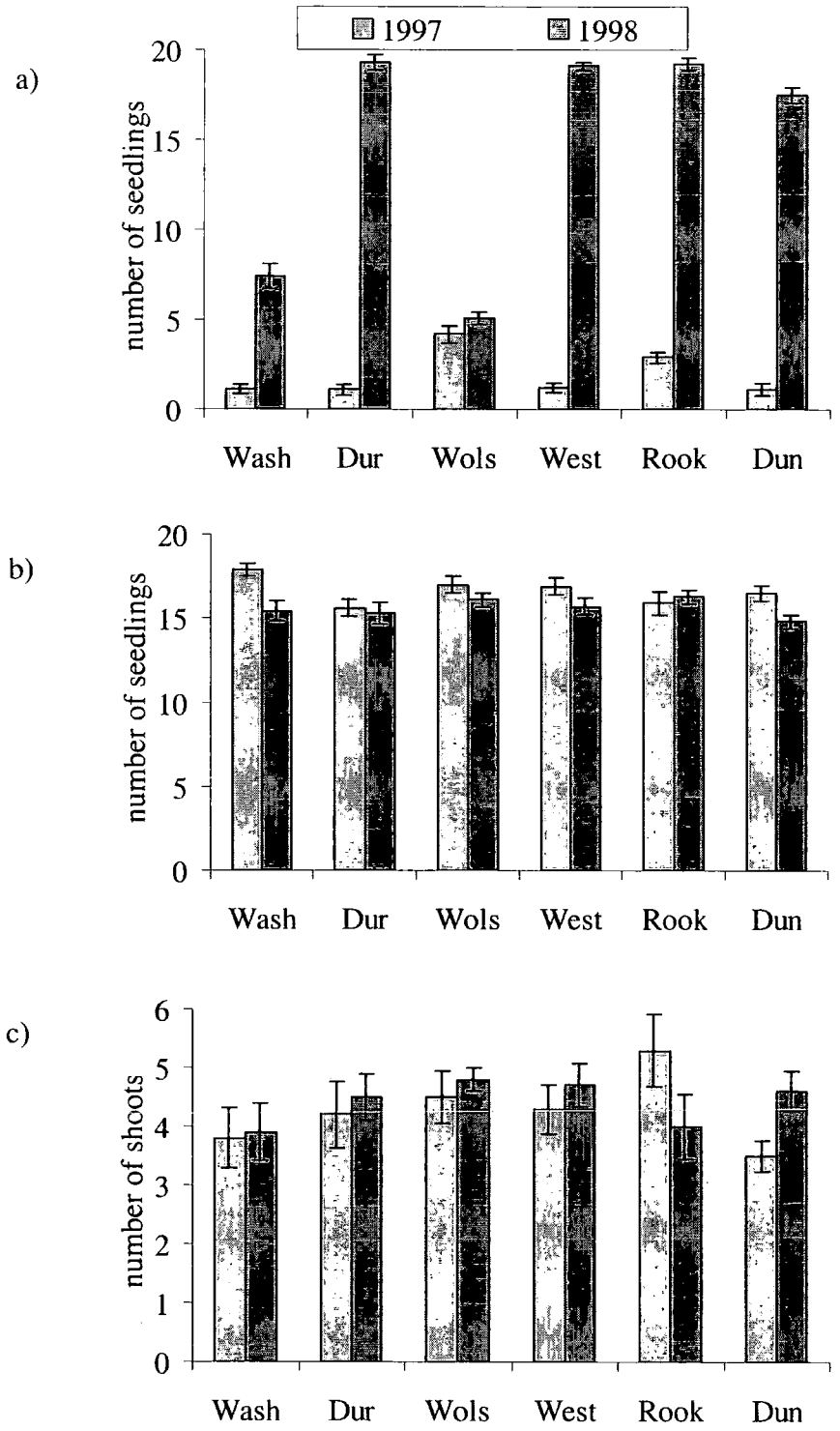

Figure 7.7 Maximum germination of a)*Impatiens* seedlings, b)*Heracleum* seedlings and c)*Fallopia* shoots produced in replicate pots across the experimental transect in 1997 and 1998. Seedlings originated from 20 seeds planted in each replicate pot in the preceding autumn and *Fallopia* shoots from five rhizome fragments.

The proportion of *Impatiens* seeds germinating was significantly different both between sites in any year and also between years ($F_{194,5}$=119.9, P<0.001 & $F_{194,1}$=3185, P<0.001 respectively). The interaction of site and year was also significant ($F_{194,5}$=174.8, P<0.001), probably due to reduced germination at Washington and Wolsingham in 1998 compared to the other sites.

There was a significant difference in *Heracleum* germination between years ($F_{194,1}$=3.45, P<0.001), with some reduction in germination in 1998 at most sites. However there was no difference in germination between sites in either year ($F_{194,5}$=11.98, NS). Conversely *Fallopia* showed no significant difference in shoot production either between sites in any year or between years ($F_{194,5}$=1.075, NS & $F_{194,1}$=0.335, NS respectively).

For all three species no trend of changing proportional "germination" in relation to altitude was apparent. The over-winter survival of *Impatiens* seeds from planting to germination was however found to be significantly related to the number of winter frost-days at a site (Figure 7.8) (r^2=0.120, $F_{118,1}$=160.44, P<0.001), though the fit of the line is poor. None of the other species demonstrated a significant relationship between over-winter propagule survival and frost incidence. Similarly no species showed any relationship between propagule survival and winter heat-sum.

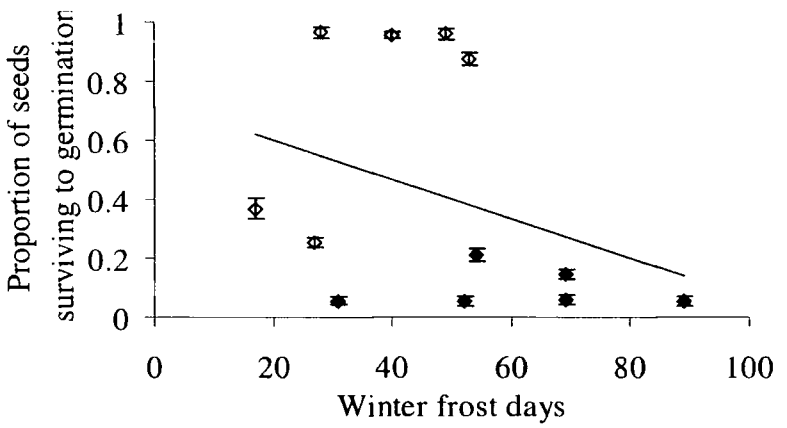

Figure 7.8 Proportion of *Impatiens* seeds surviving to germination in the spring of 1997 (filled) and 1998 (open) against the number of previous Winter frost days (Oct-Mar). Plotted values are mean values for the altitudinal sites separated into different study years.

7.3.2. Growth

The growth of the three species was measured through the growing season using plant height as a non-destructive method of estimation. To reduce any potential effects of shading in the comparison of between-site and between-year growth, only the tallest plant in each of the replicate pots was used in the analyses (Figures 7.9-7.11).

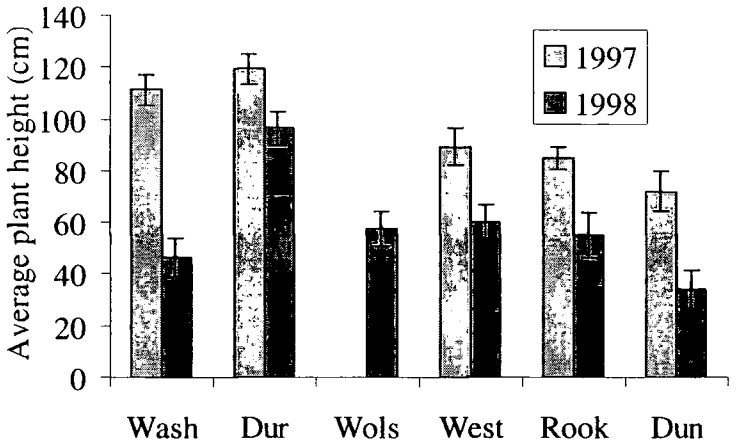

Figure 7.9 Mean maximum heights of *Impatiens* plants grown in replicate pots at the six study sites during 1997 and 1998.

Impatiens plants show a general trend of reduced height with increased altitude in both years with the exception of the most lowland site at Washington where heights are lower than Durham, which is at a marginally greater elevation. The differences in height between sites are significantly different both between years ($F_{95,1}$=88.078, P<0.001) and between sites in any year ($F_{95,4}$=16.772, P<0.001). There was also a significant interaction between year and site ($F_{95,4}$=2.774, p<0.05) arising due to inequalities of height variation between years in the different sites. The site at Wolsingham was eliminated from these analyses (and the subsequent analyses on *Fallopia* and *Heracleum*) as the plants were heavily grazed by cattle at one point.

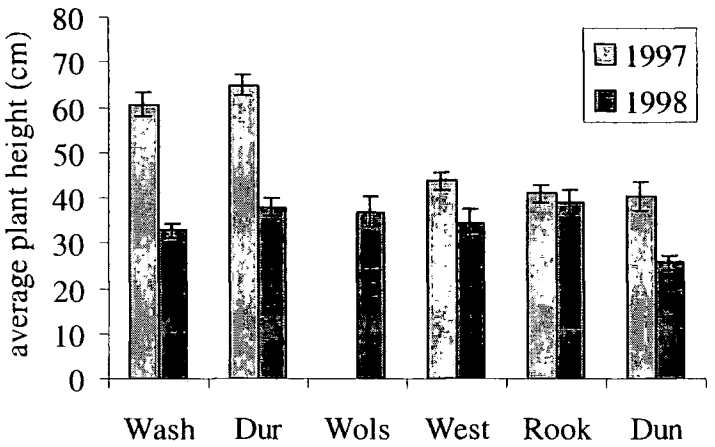

Figure 7.10 Mean maximum heights of *Fallopia* plants grown from rhizomes at the six study-sites during 1997 and 1998.

The height of *Fallopia* plants varied significantly both between sites ($F_{95,4}$=18.454, P<0.001) and also between years ($F_{95,1}$=116.476, P<0.001). In 1997 the two lowland sites had much taller plants than the upper catchment sites, with Durham producing the largest plants. However in 1998 this trend disappeared and only Dun Fell had smaller plants than the lowland sites. The lack of differentiation between the lowland and upland sites in 1998 arose from reduced stature of the lowland plants in 1998, which is probably responsible for an interaction between site and year ($F_{95,4}$=11.141, p<0.001).

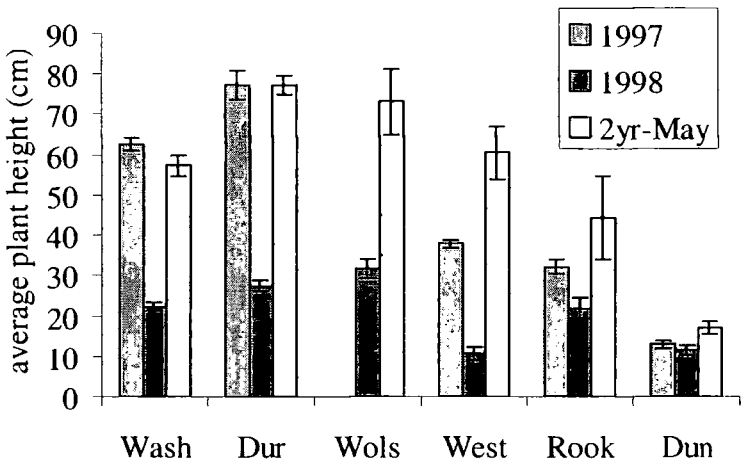

Figure 7.11 Mean heights of *Heracleum* plants in a) their first summer; labelled 1997 and 1998 to indicate the year of growth, and, b) plants in mid-May of their second summer (labelled 2yr-May). Averages are calculated for each of the six study sites.

Heracleum plant heights were measured for plants growing from seedlings in 1997 and 1998, and also for the second year growth of the seedlings from 1997 (Figure 7.11). If the growth of plants in their first summer are initially compared, it is apparent that there is height variation both between sites in any year ($F_{95,4}$=145.8, P<0.001) and also between years ($F_{95,1}$=484.1, P<0.001). There are also differences in plant heights at sites between years; for example Dun Fell shows only small height variation between years whilst Durham shows much greater variation, which gives rise to a significant interaction between site and year ($F_{95,4}$=59.0, P<0.001).

Although the second year growth is not measuring two full growing seasons (plants having to be prematurely harvested to prevent shading of experimental plots) it can still be seen that plants very quickly reach or even surpass the maximum height of the previous year, probably as a result of increased photosynthetic ability. The trends in height of these plants are similar to those noted in the other species, namely an increase in height with decreasing altitude to the Durham site followed by a small decrease in size at the more coastal Washington site.

The maximum height of all three species in their first growing season could be correlated with the summer heat-sum of sites (see Figure 7.12) and significant regression lines fitted to the data. Better regression fits using mean values for sites (not shown) suggest within-site variability of plants may mask these trends to some extent.

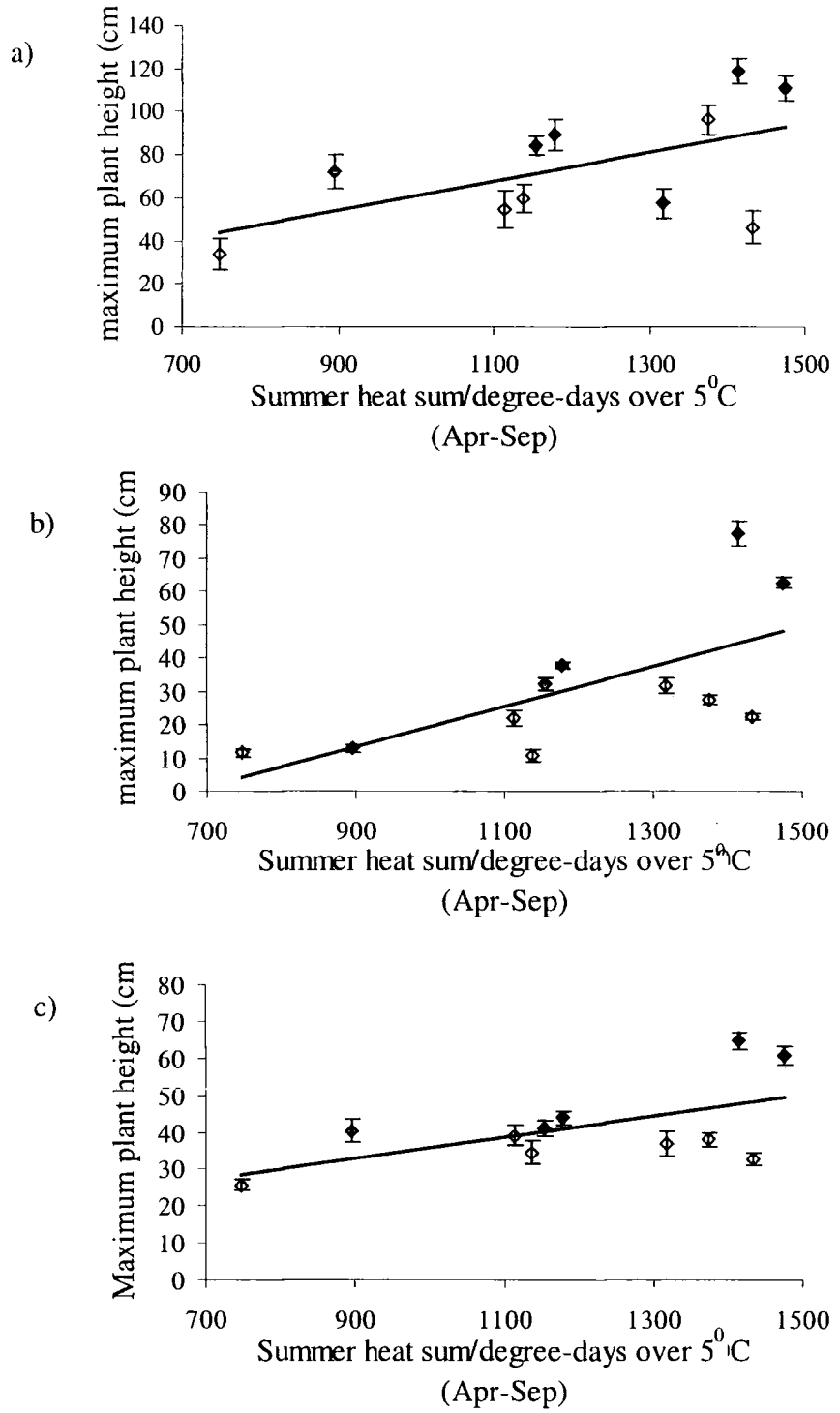

Figure 7.12 Maximum heights of a) *Impatiens* and b) *Heracleum*, and c) *Fallopia* in relation to Summer heat sum (Apr-Sep) at sites along the altitudinal experiment during 1997 (filled) and 1998 (open). Displayed height values are means (with S.E.) from each of the study plots in a particular year. Regressions: *Impatiens*: Y= 0.056x + 9.721, r^2=0.118, $F_{118,1}$=15.858. *Heracleum*: Y= 0.06x – 41.405, r^2=0.407, $F_{108,1}$=74.194. *Fallopia*: Y= 0.028x + 7.353, r^2=0.224, $F_{106,1}$=31.811. P<0.001 for all three species.

7.3.3. Survivorship

Survivorship of seedlings of *Heracleum* and *Impatiens* plants was followed at the altitudinal study sites, monitoring the survival of species at different stages of their life cycle, so as to eliminate the effects of timing differences of growth between sites. *Heracleum* seedlings in 1997 were followed up to their removal in May 1998 (Figure 7.13a), whereas those planted in 1998 were only followed to the autumn of the same year (Figure 7.13b). Similarly *Impatiens* being an annual species was only followed from germination in the spring through to the autumn of the same year.

Survivorship analyses were not undertaken for *Fallopia* plants as production of shoots continued through much of the growing season, and also because of a difficulty of ascribing visible shoots to different individual plants.

For *Heracleum* plants (Figure 7.13) germination from seeds was high in both years, with little mortality between the period of seedling appearance and the first-leaf stage. Most of the mortality in their first year occurred during the summer between the first-leaf stage and early autumn, though even in this period mortality was low. No obvious trends in survival were apparent across the altitudinal transect during this period.

Those plants which were followed into their second growth season (Fig 7.13a) showed some variation in mortality across the altitudinal gradient, both in terms of overall mortality and also in periods of mortality. For the two lowland sites (and also the site at Westgate) mortality continued at a similar low rate to that of the previous year, such that the period of greatest mortality occurred in their first year between the first-leaf stage and the autumn. By contrast the two upper sites (Rookhope and Dun Fell) showed a marked over-winter mortality, with lower mortality rates during the second summer, leading to lower plant survival than at the lowland sites. The plants at Wolsingham germinating in 1997 could not be compared to those at other sites due to a grazing incident during the first summer, which may have led to increased mortality.

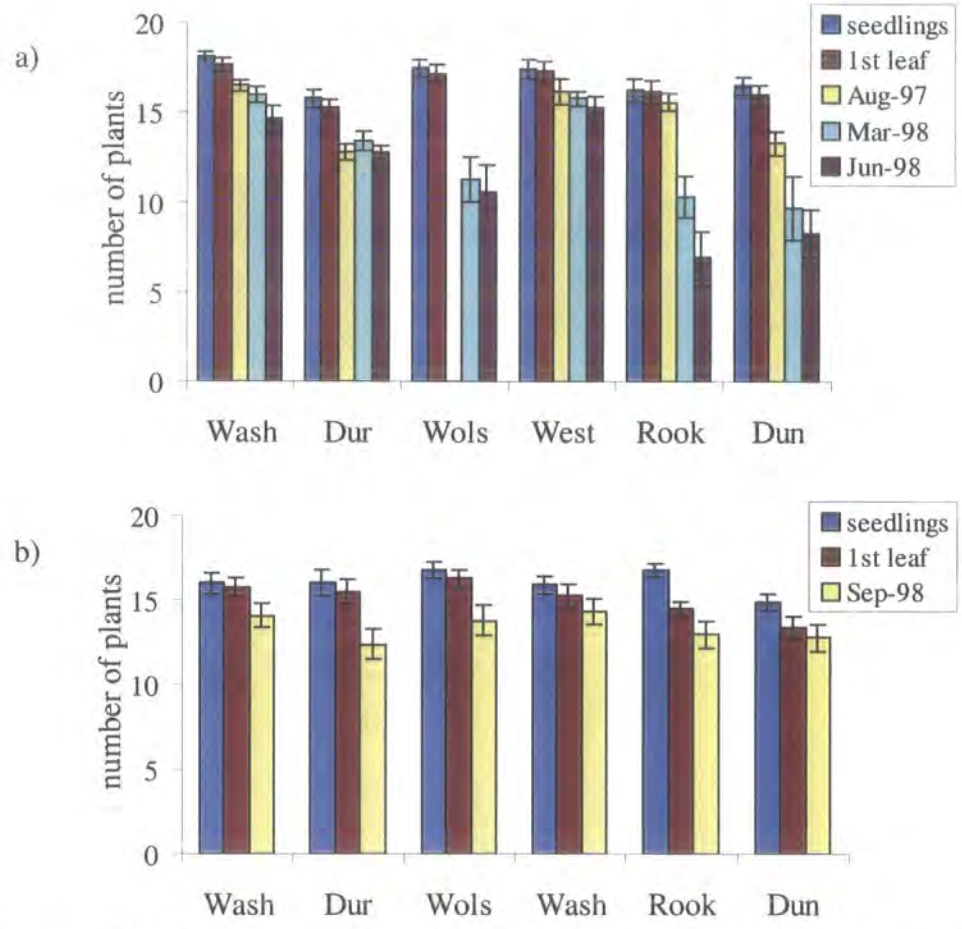

Figure 7.13 Mean survival of 20 *Heracleum* seeds across the experimental transect. a) follows seeds from autumn 1996 to June 1998; b) follows seeds from autumn 1997 to June 1998. Plant survival is shown at the seedling stage (seedlings), at the stage of producing their first true leaves (1st leaf) and at the time of harvesting (June 1998 for (a) and Sept 1998 for (b)). Two additional intervening survival stages are shown for chart (a).

Impatiens seeds showed a vastly differing germinability both between sites and also between years as shown previously. From the poor germination in 1997 there was corresponding high survival of plants at all sites through to flowering with no differences between sites (Fig 7.14). By contrast the high seedlings numbers of 1998 showed high mortality such that by the end of the year flowering plant densities were only slightly greater than in 1997. Those sites with the lowest germination rates (Washington and Wolsingham) had higher survival rates than the sites of higher germination so that the resultant flowering plant densities were similar to the other sites. The period of greatest mortality varied between sites: most of the mortality at Durham occurred after the appearance of the first true-leaf, that at Westgate and Rookhope occurred in both the seedling stage and the period after true-leaf production, while at Dun Fell most mortality occurred before true-leaves were formed. Although mortality occurred during different

phenological stages at different sites, the mortality mostly occurred in the same short time-period across the transect; hence the differing mortality of phenological groups between sites only reflect the retarded development rates of the upper sites.

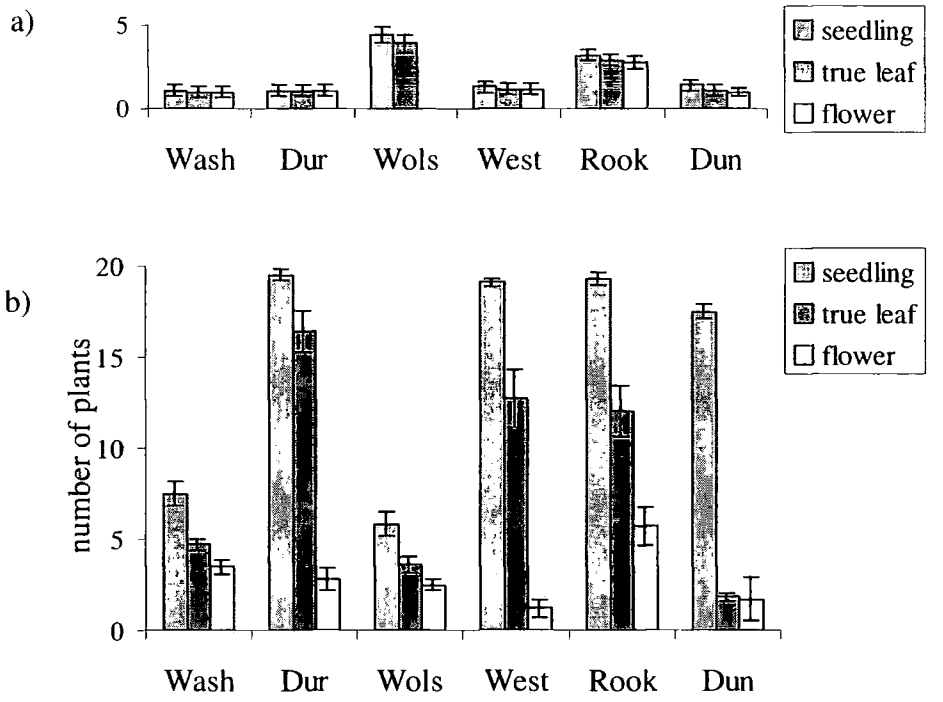

Figure 7.14 Mean survival of *Impatiens* plants along the experimental transect during a) 1997 and b) 1998. Seed were sown in the autumn preceding the respective growing season. Plant survival is shown to the seedling stage (seedling), the time of first true leaf production (true leaf) and to flowering (flower) for the two respective years.

7.3.4. Biomass

The total biomass of first year *Heracleum* plants (Figure 7.15) was significantly different between sites ($F_{783,5}$=24.949, P<0.001) with a general trend of reduced biomass towards the ends of the transect (though plants at Westgate and Rookhope buck this trend). Had plants been destructively sampled for biomass in 1997 they would have had greater biomass than the plants in 1998 (from measured maximum heights and observed size differences). From the recorded values of root and leaf biomass it would seem that there was variation in root:shoot ratio along the transect, though no consistent trend with altitude.

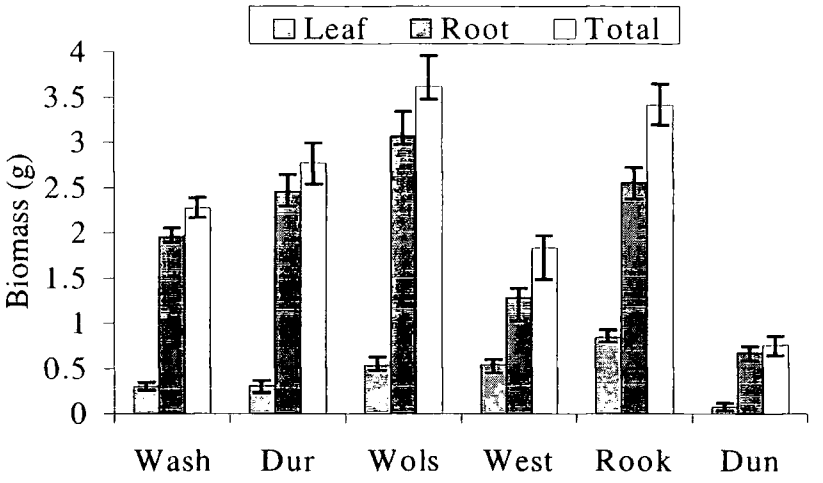

Figure 7.15 Biomass of 1st-year *Heracleum* plants grown in the study sites. Weight is partitioned into above ground biomass (leaf), below ground biomass (root) and a combination of these two (total). Plant weights were taken in September of their first growing season (1998).

The trend in biomass of harvested second year *Heracleum* plants across the altitudinal transect (Fig 7.16) are very similar to those obtained for the first year plants. There is again a significant difference in total biomass between site ($F_{50,5}$=3.855, P<0.01), with the variation due largely to reductions in growth at the extremes of the transect. The central four sites along the transect have very similar values of plant biomass and again there is no consistent trend in changing shoot:root ratio between sites.

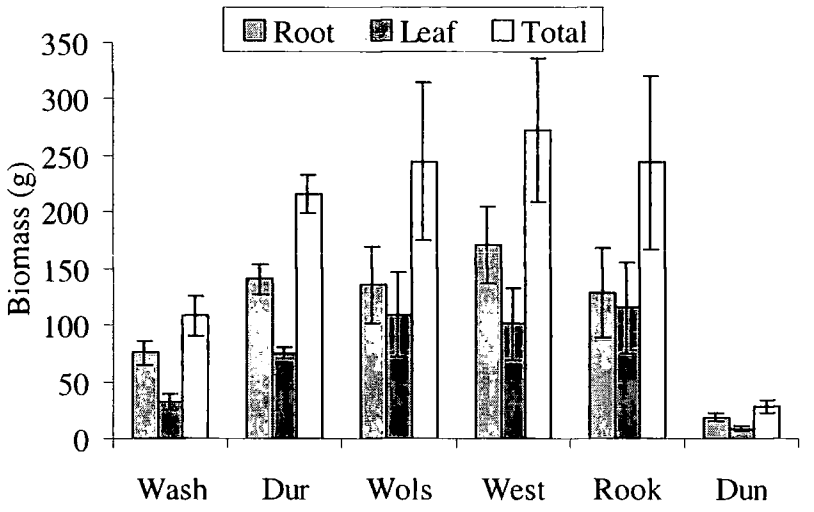

Figure 7.16 Biomass of 2-year old *Heracleum* plants grown in the study sites. Weight is partitioned into above ground biomass (leaf), below ground biomass (root) and a combination of these two (total). Plant weights were taken in June of their 2nd growth season (1988).

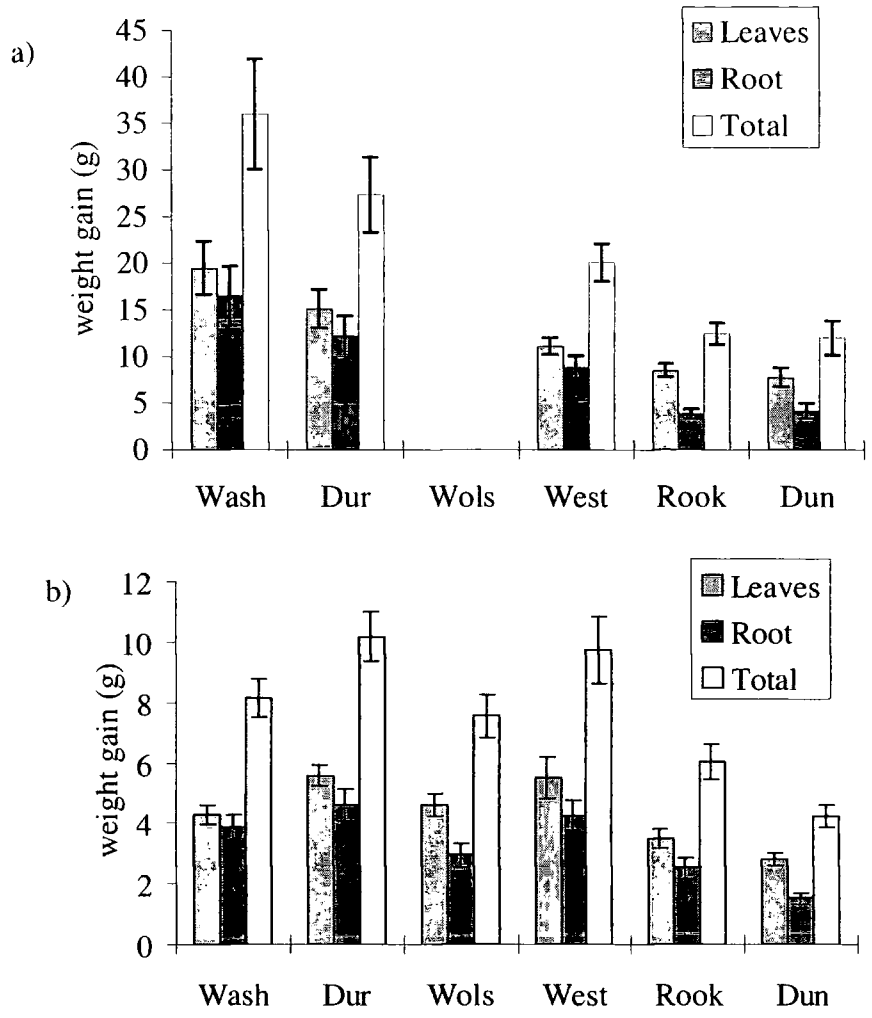

Figure 7.17 Biomass increase during a)1997 and b)1998 of *Fallopia* plants grown in the study sites. Weight is partitioned into above ground biomass (leaves), below ground biomass excluding the original rhizome fragment weight (root) and a combination of these two (total).

The biomass figures for *Fallopia* (Figure 7.17), ignoring the damaged Wolsingham site, show significant variation between site in any year ($F_{4,371}$=10.268, P<0.001) and also between years ($F_{1,371}$=128.4, P<0.001). Different patterns in the biomass of sites between years gives rise to a significant interaction between the two variables ($F_{4,371}$=7.94 P<0.001). The plants grown during 1997 show a consistent reduction in biomass with increasing altitude (with no moderating coastal effect at Washington). The most significant biomass differences between the plants grown in 1997 and those grown in 1998 is the reduction in biomass accrued during the growing season. Plant weights during 1998 are in the region of 25-50% of those recorded in 1997. A decrease in biomass with increased altitude is also less apparent in 1998, only the uppermost sites showing reduced biomass compared to those at lower altitude.

There is a slight trend of *Fallopia* plants at the upland sites partitioning relatively greater proportions of their biomass into above ground growth compared to the lowland sites.

The biomass of first-year *Heracleum* plants grown during 1998 was significantly related to the summer heat-sum for that year (r^2=0.059; $F_{792,1}$=49.587, P<0.001, Figure 7.18). However with only six values for heat-sum the regression is heavily dependent upon the low weights of plants at Dun Fell (bottom left point on figure) to produce a significant fit. A similar exercise undertaken on the harvested two-year old *Heracleum* plants could detect no linear relationship between plant biomass and the heat-sums of the two summers, either separately or treated together. However if the totalled summer heat-sum for the two years (up to the point of harvest) is compared to biomass using a quadratic regression equation, a good fit is obtained (Figure 7.19). This would seem to indicate that the growth of *Heracleum* increases with summer temperature up to a point (about 1700 degree-days), after which plant productivity is reduced. It may however be that other factors (e.g. low precipitation) are affecting growth at these lower sites.

There is much variability of two year-old *Heracleum* plants within sites, such that using mean values at sites to fit the quadratic equation produces a better r^2 value.

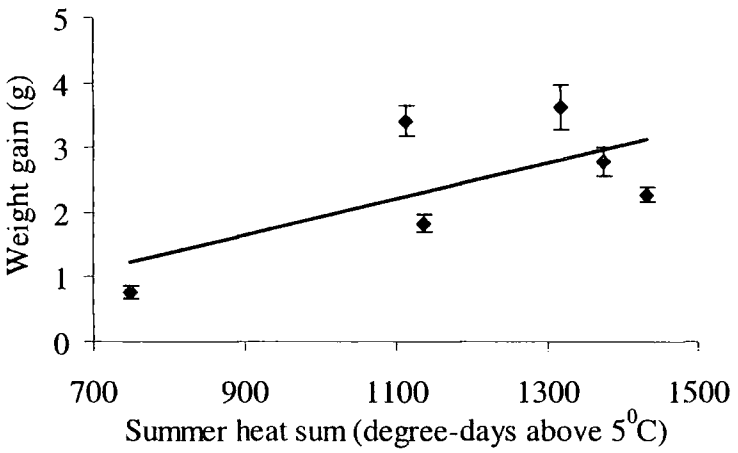

Figure 7.18 Mean weight gains of *Heracleum* plants along the altitudinal transect in relation to summer heat-sum. Plants were grown from seeds sown in 1997 and were destructively sampled in the early autumn of 1998.

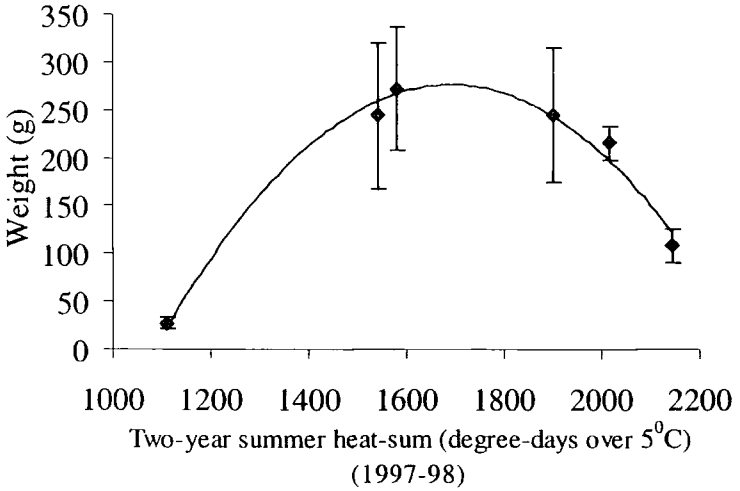

Figure 7.19 Biomass of two year-old *Heracleum* plants along the altitudinal transect in relation to summer heat-sum over the period of growth. Plants were grown from seeds sown in 1996 and were destructively sampled in late May 1998. Regression: r^2=0.3029, $F_{55,1}$=9.951 P<0.001, $Y=-0.0008x^2 + 2.7041x - 1987.3$ Using means only r^2=0.9815, $F_{5,1}$=79.385 P<0.01. $Y=-0.0007x^2 + 2.536x - 1869.1$

Weight gain by *Fallopia* plants during their first year is correlated with the summer heat-sum of that year (r^2=0.086; $F_{416,1}$=39.262, P<0.001, Fig 7.20). Again there is variation in plant biomass that cannot be explained using only the heat-sum variable.

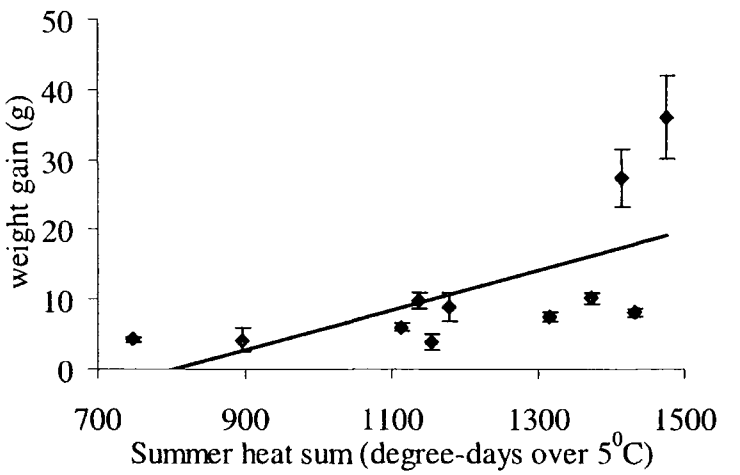

Figure 7.20 Weight gain of *Fallopia* plants grown along the altitudinal transect in relation to summer heat-sum at the sites over the two study years. Plants were grown from rhizome fragments and harvested in the early autumn.

7.3.5 Flowering & seed production

a) Timing of seed set

An earlier date of seed development was found for *Impatiens* at the lower altitude sites (Figure 7.21) and, in 1998 at least, plants in the upland areas were unable to ripen seeds

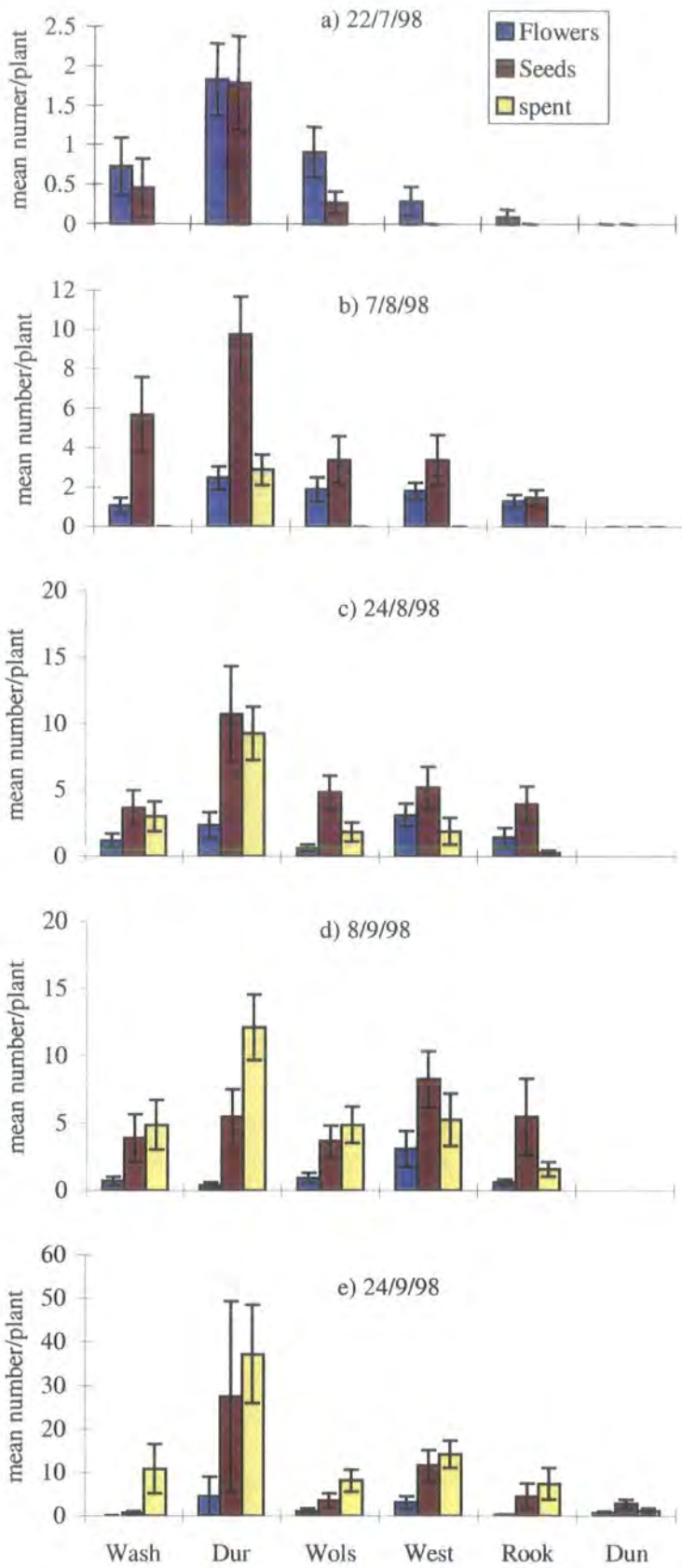

Figure 7.21 Flowers, seed-pods and spent seed-pods on *Impatiens* plants in the altitudinal experiment at different dates during the flowering period.

in any number before the end of September. Even the lowland sites produced a good proportion of their seeds after the start of September. It can be seen that the number of flowers on plants at any time is fairly consistent across the transect and through much of the flowering period, only dropping in number after September. The production, initially of ripening pods, and subsequently spent pods can be seen to follow a pattern, with largest numbers at the Durham site and reduced numbers as altitude increases and towards the coast at Washington.

b) Seed output

Flowering only occurred on one *Heracleum* plant along the experimental gradient. This occurred at the Wolsingham site in the middle of the transect. Successful pollination occurred and seeds were set. The umbels produced and the plant stature in general were much less than in observed natural plants (height of 1.52m compared with 3-4m in natural growing plants), presumably due to being pot-bound. An estimated potential seed output (see Table 7.2) of c12,700 compares to a realistic value in the region of 3-4000, due to the failure of the satellite umbels to produce seed, presumably due to a lack of resources.

Table 7.2 Description of seed-head characteristics of the flowering *Heracleum* plant at Wolsingham.

Umbel type	No. of Umbels	Rays on umbel	Florets/ray	Seed estimate	Total seed estimate
Primary	1	52	43.7 $\bar{+}$2.45	4,545	12,712
Satellite	4	41.5 $\bar{+}$3.38	24.6 $\bar{+}$1.44	8,167	

As only one inflorescence was produced at the altitudinal sites it was not possible to examine the effect of altitude on the productivity of this species from this experiment. However as an alternative brief test, plants were sampled at their upper limit along the Durham rivers (on the Tees) and their fecundity compared to that of plants growing in a similar habitat in the lowlands. This data indicated that fecundity was no different between the two sites (t-test: n=20, t = 0.598, N.S.).

The seed outputs of *Impatiens* plants across the altitudinal transect were inferred by measuring the production of seed-pods along with the seed content of such pods. Other factors such as seed weight and changes in seed content of pods over time and from different areas of the plant were also recorded

The mean numbers of pods produced per plant across the transect (Figure 7.22) show

significant variation both between sites ($F_{136,4}$=9.264, P<0.001) and also between years ($F_{136,1}$=27.504 P<0.001). The patterns of pod production between years at sites gives rise to a significant interaction between the two variables ($F_{136,4}$=9.026, P<0.001). The pattern of decreasing pod production with altitude observed in 1997 is not apparent in 1998, although if the two aberrant sites (in terms of germination success and date) of Washington and Wolsingham are overlooked the general trend remains. All of the sites produced significantly lower pod yields in 1998 than in 1997, this trend being most pronounced at the Washington site.

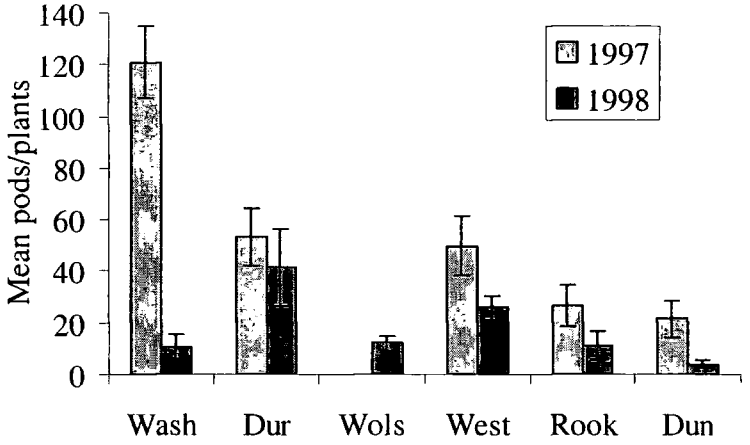

Figure 7.22 Mean seed pod production of *Impatiens* plants grown at the six study sites during 1997 and 1998

The pod production figures per plant are compared to summer heat-sum values over the period of plant growth in Figure 7.23. This indicates a relationship between pod output and heat-sum, higher summer heat-sum generally resulting in increased pod production per plant (r^2=0.074; $F_{161,1}$=12.802, P<0.001). The large amount of variation about the regression line indicates that heat-sum is not the only factor affecting pod production.

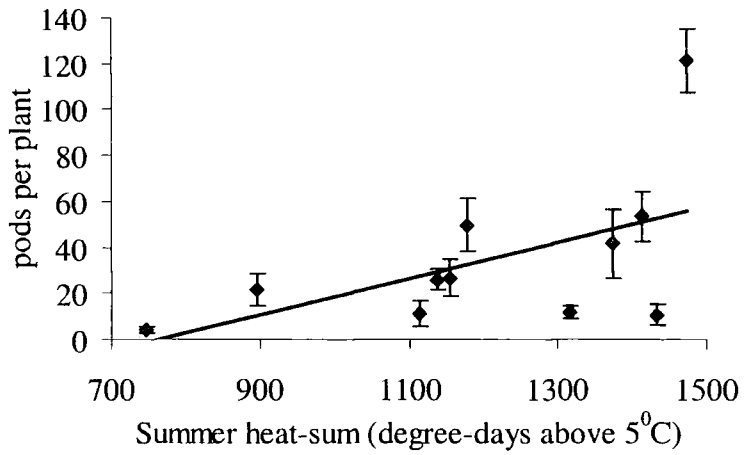

Figure 7.23 Pod production of *Impatiens* plants in relation to summer heat-sum (Apr-Sep) at sites along the altitudinal experiment over the two years of study. Displayed values are means (with S.E.) from each of the study plots in a particular year.

The patterns of seed content per pod, monitored during the 1998 growing-season, indicated that there was variation both between sites and also within sites at different times of the year (Figure 7.24). Analysis of pod content between sites irrespective of sampling date confirmed that seed content varied between sites ($F_{761,5}$=45.436, P<0.001). All sites with greater than two dates of seed collection showed a trend of initially high seed content of pods followed by a decrease in seed numbers and a subsequent partial recovery. This trend is most obvious for the Durham site where there is a continuous sampling record. A further analysis of the seed output of an individual plant at the Durham site (Figure 7.25) highlighted the fact that during the period of pod production the number of seeds per pod from any inflorescence decreases as the growing season progresses. Additionally it was noted that the seed production by secondary inflorescences (i.e. those growing from side branches of the main stem) began later than for the primary inflorescences but followed the same pattern of seed decline per pod. Significant regression lines could be fitted to the plots of seed content of pods over time (r^2 =0.535; $F_{166,1}$=190.8, P<0.001 for primary inflorescences: r^2=0.251; $F_{830,1}$=229.2, P<0.001 for secondary). Also, because of the delayed onset of seed production on secondary inflorescences, at any given time there were more seeds in pods from these latter inflorescences than from primary ones. This may explain the partial recovery in pod sizes observed towards the end of the growing season; the primary pods having all been set leaving only the secondary pods (with their greater seed content) remaining.

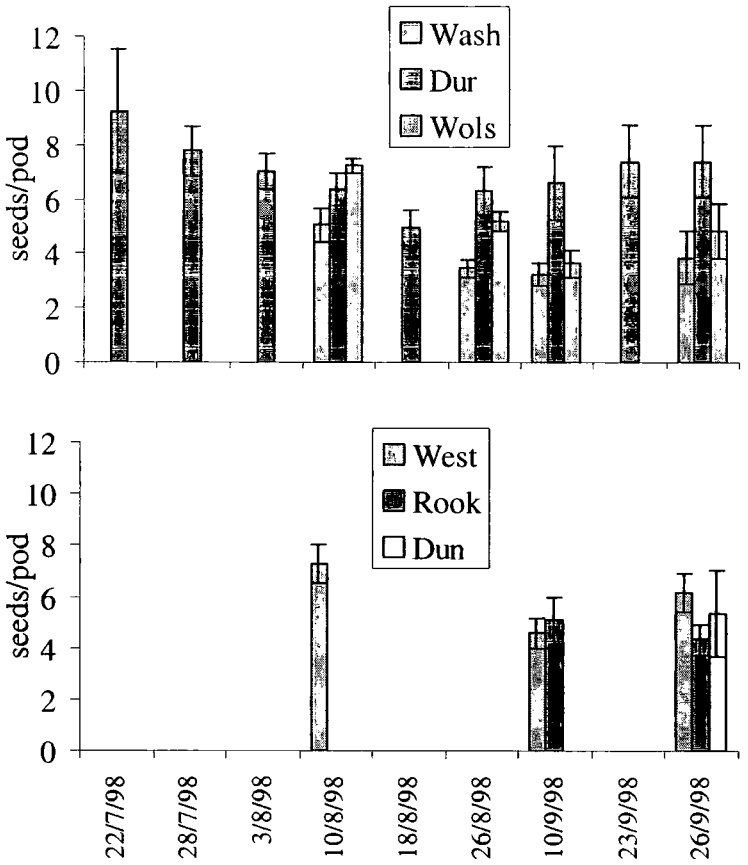

Figure 7.24 Changes in seed content over time of *Impatiens* pods from plants grown at sites along the altitudinal transect. Mean values are calculated from median pod size of each replicate plant to standardise for varying pod output per plant. Part a) depicts results for the three lowland sites and part b) shows similar results from the upland sites.

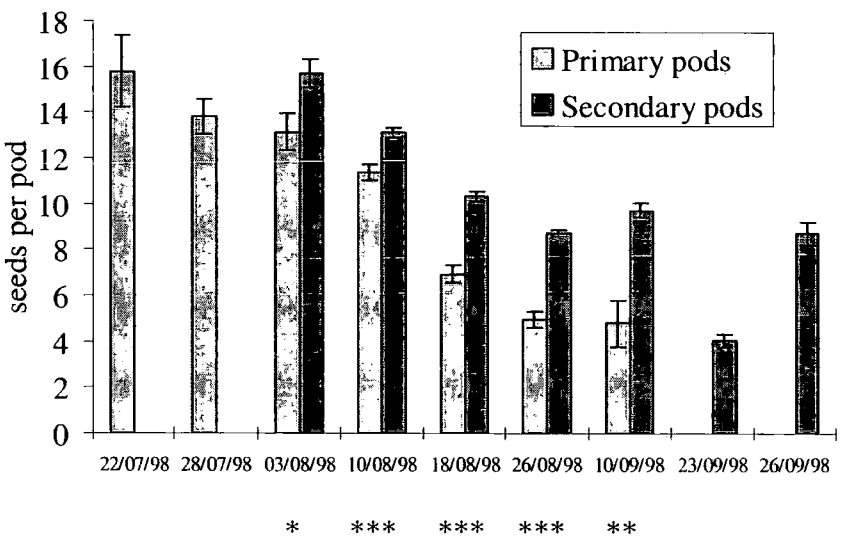

Figure 7.25 Seed content of *Impatiens* pods from a single large plant. Pods are separated into those from the primary and secondary inflorescences and by the date of sampling. Differences in seed content of pods from different inflorescences on any date were tested using t-tests; significant differences are marked *, ** & *** for levels of 0.05, 0.01 and 0.001 respectively.

c) *Impatiens* seed weight

Seed masses were recorded along the altitudinal gradient and compared, both within sites at different times and also amongst sites. (Figure 7.26). These data show a trend of increasing weight per seed over time at the lowland sites, a trend that disappears at the higher altitude sites. An analysis of seed weights at the various sites (irrespective of time) also shows a significant difference in the mean seed weights at the sites ($F_{1684,5}$=74.50, P<0.001). Comparing the seed weights from the earliest time of seed production at each of the sites (i.e. standardising the morphological growth stage) also indicated significant differences ($F_{134,5}$=24.724, P<0.001).

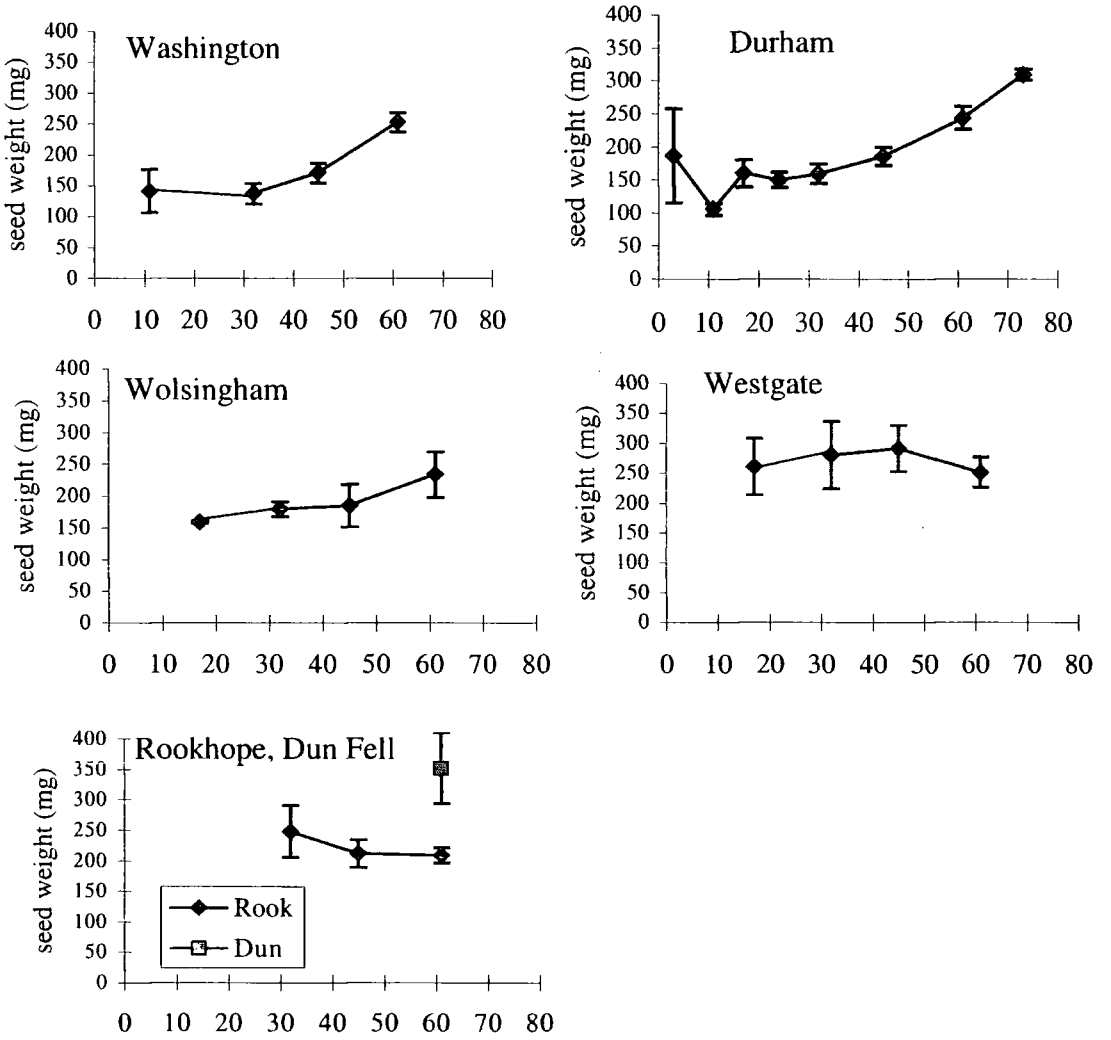

Figure 7.26 Change in mean seed weights at the altitudinal study sites against time (in days) from first seed outputs (22nd July).

A comparison between sites shows that the upland sites had large seed weights, whereas the lower altitude sites initially produced lighter seeds but seed weight increased as the

season progresses such that, by the end of the season weights were comparable to the upland sites.

By comparing the mean seed weights from plants at the various sites (over the entire growing season) to climatic variables an idea of the mechanisms causing seed weight variation can be gained. The data collected from the altitudinal sites through the growing seasons showed that the usual time taken for a plant to produce a flower, ripen the pod and set seed was approximately four weeks. To establish whether the climate during this period was related to the final weight of spent seeds, mean seed weights were plotted against the heat-sum during the preceding four weeks and a regression line fitted to the data (Figure 7.27). This analysis demonstrates a strong relationship between the heat-sum during the time of pod development and the weight of seeds produced (r^2=0.227, $F_{1685,1}$=493.48, P<0.001). Whether the increase in seed weight during cooler periods arose due to a reduced seed content of pods during this time could not be reliably examined, as seed were collected in such a way that weighed seeds could not be associated with their corresponding pod content.

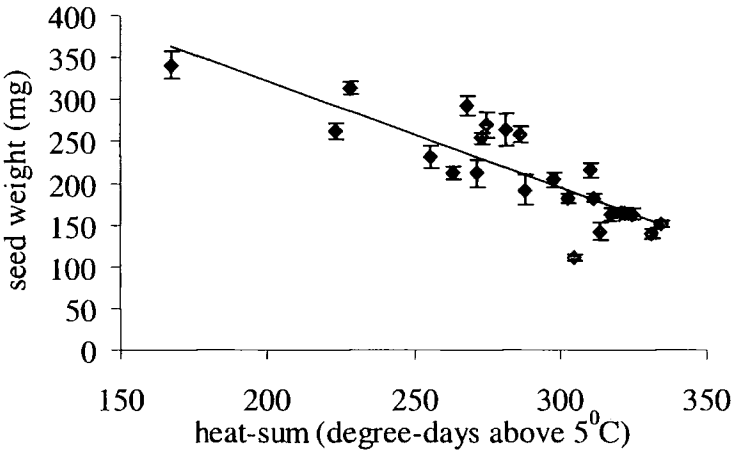

Figure 7.27 Weight of seeds from *Impatiens* plants grown at sites along the altitudinal transect in relation to heat-sum over the preceeding four weeks. Displayed values are means (with S.E.) from each of the study plots separated into different collection dates.

7.3.6 Herbivory

The survival of seedlings of *Impatiens* and *Heracleum* was found at several sites to be influenced by slug predation, the slugs operating by chewing through the stems of the seedlings at ground level and effectively felling them. It was also found that many of the muslin bags containing ripe *Impatiens* seeds had been predated by what was assumed to be a small rodent species. The bags particularly affected were those which had fallen off

the plant onto the ground, although some bags had been predated whilst still on the plants.

At the lower elevation sites, *Impatiens* seedlings often suffered herbivory to varying degrees. This could probably again be attributed to slugs or some similar predator. There was also some herbivory of the leaves of *Fallopia japonica* at one or two of the intermediate altitude sites during the summer when the plants were quite well grown. This damage was attributable to the beetle *Gastrophysa viridula* and generally took the form of missing segments of leaf. However the latter damage was probably not sufficient to impair the growth of the plants significantly.

7.4 Discussion

Much work has been undertaken on the effects of climate on plant species distribution (Woodward 1987 provides a good summary) and work has been published previously on the effects of climate on the occurrence of the study species (e.g. Beerling *et al.* 1995; Beerling 1993, 1994). However few works have examined the effects of climate on these species in terms of growth and productivity traits.

The large variation in germination dates noted for all the species both between sites and years was linked to the number of winter frost days in the spring, increased frost incidence delaying germination date, and also to spring growing degree-days (over 5^0C). For all three species spring heat-sum proved better than frost-days in explaining germination dates, suggesting that this was the more important determinant, though periods of low temperature are necessary to break seed dormancy (Shumova 1972, Grime *et al.*1981).

The mean dates of germination of *Heracleum* seed during 1998 are comparable to those mentioned in the literature (e.g. Jan-Mar, Tiley *et al.* 1996) whereas the germination dates noted at the sites in 1997 (through May) are unusually late. Similarly with *Impatiens* the germination dates in 1998 for all sites except Dun Fell fall into the time of reported germination (Feb-Mar, Beerling and Perrins 1993) whereas those in 1997 are considerably later. A comparison between plants germinating along the Wear catchment and those in the experimental pots indicated that the pot grown plants had a delayed germination date of 1-2 weeks compared to plants in natural situations in both years. It is possible that the germination dates of both species were slightly retarded in the experimental plots as a result of not planting seeds out until early winter (late December

in 1997) so that they miss some of the climatic hardening and dormancy breaking conditions of late autumn.

The climatic temperature anomalies for the two winters will also partly explain differing germination dates; 1998 having warmer spring temperatures and a milder preceding winter than did 1997.

Mean dates of shoot production of *Fallopia* are later than those given by Beerling *et al.* (1994) (early March to late April). This is probably due to recording shoot appearance through the year in the experiment and not spring "germination" *per se*.

The very high germination rates of *Heracleum* across the transect were surprising given the species current distributional limit in the county. Similarly the high germination of *Impatiens* seeds (in 1998 only) and *Fallopia* rhizomes in the upper sites was contrary to expectations based on their current distributions. It seems that winter temperatures are not sufficiently low across the transect to restrict the occurrence of any of the species through over-winter seed/rhizome mortality; the general lack of association between winter frost days and mortality for all species except *Impatiens* supports this. The substantial germination rate of *Heracleum* seeds at higher altitudes is concurrent with the species native habitat, where cold winters occur (Mandenova 1950, in Tiley *et al* 1994). This suggests that the current lowland distribution of *Heracleum* in the UK may be due to dispersal limitations and human influences, as in the Czech Republic (Pyšek 1994).

The contrast between the very high germination rates of *Impatiens* seeds at some sites in 1998 and the low rate at all sites during 1997 and at Washington and Wolsingham in 1998 is perplexing. Although winter frost-day sums were shown to significantly affect seed germination success the four values of high germination still lie as outliers on the regression slope and it seems that factors other than climate also served to affect germination success. Examination of the climatic conditions of the lowland sites during 1998 revealed that there were no differences in the spring climate variables (e.g. late frosts) which could explain the differential germination. It is also possible that the different planting dates and seed storage times affected seed mortality between years but again this does not explain the low germination at sites in 1998.

Seeds originating from different plants may have had varying degrees of viability but this also doesn't explain the differences between sites in 1998 where all seeds originated from the same general seed stock. Other explanations that can be offered but which

cannot be confirmed include slight variability in storage/transport of seeds planted at different sites, *in-situ* seed predation or attacks from soil based pathogens or viruses (as found by Prowse, (1998)). It is perhaps worth quoting Rorison (1981) with respect to these inexplicable variations, who stated

> "A suitable experiment...(to follow seedling establishment and survivorship)...is easy to plan but not to execute. The vagaries of climate and predation can play havoc with replicated plots in the field"

The high over-winter survival rates of *Fallopia* rhizome fragments even at the upper sites also seemed surprising until climate data were compared to the upper limit of the species in its native Japan and related seedling survival. Here the winter climate is much cooler than the climate the plants experienced on Dun Fell (using climate data for Japan from the IIASA database, Leemans and Cramer 1991). Maruta (1983) found 100% survival of seedlings on Mt Fuji with weights comparable to fresh weights less than 1g.

The maximum heights attained by the *Heracleum* plants grown in the experimental pots was probably slightly less than if grown in a non-restricted environment. The general trend for optimal height and biomass of plants to occur at intermediate sites along the transect, with reduced stature plants occurring towards the ends of the transect, suggests climatic restrictions to productivity.

Biomass values for *Heracleum* plants are low compared to plants grown from the wild (Tiley *et al.* 1996). A combination of restricted root spread and high competition probably being the major cause of this variation. Tiley *et al.* (1996) state that plants are smaller where rooting is restricted. Wild plants also partition a greater portion of biomass to their roots (Tiley *et al.* 1996) compared to the pot plants.

Although linear regressions can be fitted to the maximum heights of plants in their first year, and also to the weight gain of first-year *Heracleum* and *Fallopia* plants, in relation to summer heat sums, there is still variation about these fitted curves which suggests other factors are also important. Beerling and Perrins (1993) suggest soil moisture as controlling small vs. large *Impatiens* phenotypes, whilst Koenies and Glavec (1979) suggest growth is directly proportional to soil water.

Similarly, although germination success can be adequately predicted from spring frost days and heat sum it is likely that other factors, perhaps including short duration climatic events, will also determine germination probability. Huntley *et al.* (1998) detected plant community changes that were attributed to local changes undetectable

from mean figures. With a limited dataset of climate variables available it is likely that much of the remaining variation will also be accounted for by alternative climatic factors. It is likely that precipitation and windspeed are amongst those factors important in explaining more of the variation at the study sites. Several reports have suggested that motion *per se* is a potent inhibitor of growth (Grace 1981) and wind will also serve to drive down ambient temperatures during the growing season. Hence the increased wind-speed at the extremes of the transect during much of the year may serve to explain the observed patterns of reduced plant performance at these sites. Indeed Beerling *et al.* (1994) suggest that shelter from wind damage could be a factor determining the altitudinal limit of *Fallopia* in some areas. Similarly the reduced precipitation at the lowest sites during mid-summer may have deleterious effects on the growth of the three species, all of which have been noted to be drought intolereant to varying extents (Beerling *et al* 1994; Beerling & Perrins 1993; Tiley *et al.* 1996).

A combination of increased spring and reduced summer precipitation in 1998 compared to 1997, combined with higher summer windspeeds in 1998 (Table 7.3) may also partly explain altered productivity of the species between the two years.

Table 7.3 Windspeed and precipitation data for 1997 and 1998. Derived from the Durham observatory.

	Wind speed (summed daily mean/knots)		Precipitation (mm)	
Year	Mar-May	Jun-Aug	Mar-May	Jun-Aug
1997	227	134	77.2	305
1998	270	383	282.1	195.3

Beerling *et al.* (1994) suggest that *Fallopia* will not grow where precipitation is less than 500mm per annum, a figure lower than experienced at any of the study sites, though values of c630mm close to Washington may be low enough to restrict growth. Annual rainfall at Durham for both years was above the suggested minimum threshold. The differing growth between the two years may suggest that a summer-only precipitation limit would be more appropriate.

Unlike the previous examples it was not possible to fit linear regressions to the plots of 2 year-old *Heracleum* plant biomass against summer heat-sum. It was however possible to fit a quadratic regression to the plant weights when plotted against the cumulative heat-sums for the two seasons. This indicates that although heat-sum is obviously of some importance in determining the productivity of this species either too much or too little may be detrimental to growth. There may however also be other factors operating

at the extremes of transect. The fitted regression line suggests an optimal heat-sum value of approximately 1700 degree-days over a similar period. Such an area would lie between the sites at Westgate and Wolsingham i.e. a region far upstream of the species current distributional limit. These biomass findings fit well with the observed flowering of *Heracleum* at Wolsingham, suggesting an optimal climate.

It could be hypothesised that at lowland sites the species will be restricted to poorly drained areas due to sub-optimal precipitation but that at upper sites with warm climates and higher precipitation the species may be able to prosper away from water courses.

Survivorship values of *Heracleum* seedlings indicate high survival rates over the whole period. The studies of plants grown in natural situations, shown in the previous chapter, indicate exceptionally high mortality of seedlings. It seems that in the absence of competition (other than with siblings) the seedlings are capable of high survival. The large over-winter mortality at the upland sites (relative to the remaining sites) could also be expected given the harsher climate and the reduced rhizome reserves stored per plant. The reduced biomass of plants at the Washington site noted previously does not have a corresponding trend in reduced plant survival, seedlings surviving as well here as at any of the other sites. It seems that while factors operate to limit growth of the species at this coastal site, they do not do so to a point causing plant mortality.

The trends of *Impatiens* seedlings survival are difficult to explain. Survival of seedlings during 1997 was universally very high after the initial almost complete mortality of seeds before emergence. The two sites with low initial germination during 1998 also show only moderate mortality, whilst the other sites show high mortality between germination and flowering. The fact that the mortality at the sites of high germination during 1998 occurred at a similar time across sites irrespective of plant morphological stage indicates that intra-specific competition was not the cause. The seedlings rapidly changed in appearance from looking quite healthy to decaying remnants; the exact cause of the ailment remaining unknown. It is possible that the recorded differences in emergence between sites and years may simply reflect the same ailment affecting plants at a stage before seedlings broke the soil surface. Irrespective of the reasons for mortality of *Impatiens* plants, the resultant survival of plants to flowering was similar in both years and between all sites. This fact, coupled with the lack of an effect of altitude on germination suggests that the species is not restricted in its present distribution due to either of these factors.

The results of monitoring the timing of flowering and seed production by *Impatiens* indicate that much of the seed output occurs at the very end of the growing season at the more marginal sites. It would only require a small reduction in length of the growing season at the upper two sites to reduce seed output to almost nothing.

The different trends of *Impatiens* pod production across the altitudinal transect in the two different years can be partially explained by differing summer heat-sum values between the two years. However there are anomalies such as the contrasting high and low production at Washington in different years of similar heat-sum, suggesting that temperature alone is not driving pod production. As mentioned previously with *Fallopia*, these differences may be partly attributable to other climatic variables. Beerling (1990a) similarly demonstrated a ten-fold difference in biomass values of *Impatiens* plants between two different years in similar sites.

The analyses of seed content of pods indicates that as well as producing fewer pods the upland sites may have less seeds per pod than more favoured areas. Additionally the lowland sites produce seeds over a wider time interval, which may serve to safeguard against any catastrophes that could occur if the seeds are shed only at the end of the growing season. Timing of seed production may also have implications for the dispersal of seeds due to flood event timing, Wadsworth *et al.* (submitted) finding some periods of the year better than others for long-distance riverine dispersal.

The reason for differing seed content of pods between sites and over time could be due to a variety of factors. Just as seed weight is shown to be correlated to heat-sum of the previous month, so pod content could be determined by climate at a particular stage of development. Climate could affect pod size by causing differing growth of the pollen tube (*cf* Pigott & Huntley 1981) or some other part of the reproductive organs. Similarly it could affect pollination effort by limiting or encouraging insect pollinators. Fisher and Mallhies (1998) suggest that in *Gentianella germanica* smaller populations produce few seed-pods possibly as a result of pollinator limitation. It could be argued that a similar result might be expected at the upland sites here, where pollinators are likely to be present in lower numbers and will not be adapted to visit *Impatiens* plants.

The pod size may to some extent be pre-determined; the differing pod size between primary and secondary inflorescences may reflect this, or may be due to available resources, which in turn are related to photosynthetic capabilities. There are also likely to be genetic influences on pod production by individual plants.

Towards the end of the growing season seed weights may be larger due to unfavourable conditions preventing further pollination, thereby halting pod production and allowing more resources to be partitioned to the currently developing pods. It has been suggested for some species that seed size may vary if there is a clear distinction between conditions operating when the seed number is determined and when they are being filled (Harper *et al.*1970).

The difference in the weight of *Impatiens* seed produced, both between sites and within sites over time, could have important implications for the spread of the species should differing seed size prove advantageous. It may be that in producing heavier seeds upland sites give their seedlings more of an energy store, which could prove beneficial during the early year when growing conditions are sub-optimal. Eriksson (1999), working with *Convallaria majalis*, found increased survival of seedlings originating from larger seeds. Conversely if upland plants could redirect their resources into producing more smaller seeds this could compensate for low pod productivity and seed loss events. Determining whether such patterns of seed production are fortuitous pre-adaptations or are simply serving to hinder further spread requires additional work.

It has been shown in the case of two *Rumex* sp. that the heaviest seeds are the last to be shed, germinate later and produce larger cotyledons (Cavers and Harper 1966). If the same were true for *Impatiens* the heavier seeds produced at higher altitudes could prove to be beneficial by germinating later, thereby avoiding early frost damage, and also by carrying greater resources to oversee the cool periods characteristic of the uplands in early spring.

From the collated data it seems that the factor most likely to limit *Impatiens* at the upper altitudes of the transect is fecundity, which in turn is related to the summer heat-sum. Although the plants proved capable of setting seed at all sites it seems that at the upper sites the seed output is unlikely to be sufficient to allow for over-winter seed mortality and growing season mortality as well as natural factors such as inter-specific competition and seed loss through herbivory, flooding etc. Given the species annual nature, it would require a reasonable seed set in most years to become permanently established at any site. This combination of factors probably restricts the occurrence of *Impatiens* and allows for only periodic spread to higher altitudes during favourable seasons between which it will go extinct. The flash-floods which are prone to occur in upper catchment areas may also cause populations of the species to be more mobile and

temporary than in the lowlands where continual influxes of seeds from upstream facilitate persistence.

With *Heracleum* it seems likely that the species would be able to germinate and grow quite readily at all but the most upland site. Similarly, recorded seed output, although measured over a smaller altitudinal range, would also be unlikely to restrict the species. It may be that at sites above those at which *Heracleum* currently occurs, the growing season would not be of a sufficient length to allow seed maturation (Tiley *et al.* 1996 state that plants flowering in September may fail to set seed) but further work would be necessary to determine if this is so. At the most upland site at Dun Fell the rate of biomass production would probably be insufficient to permit flowering in a three to four year period, though again this would require quantification.

The ability of plant species transplanted outside of their range to produce viable seed has been reported for other species (e.g. Prince & Carter 1985), the actual limits being ascribed to secondary effects on populations and metapopulations rather than on individual plants. Such an explanation could equally be offered for the study species here. Prince & Carter detected an increased allocation of resources to leafy matter in their study plants (*Lactus serriola*) when transplanted beyond their natural range, a result mirrored in the *Fallopia* plants in this study. Several authors studying species towards their distributional limit have observed belated flowering (e.g. Pigott 1970; Pigott & Huntley 1981; Prince & Carter 1985), as was noted with *Impatiens* along the transect. The potential implications of this for *Impatiens* have been discussed previously. Delayed vernalisation and subsequent germination of seeds, finally leading to complete failure has also been suggested as a follow-on effect resulting from delayed flowering (Marshall 1968). This cannot be ruled out with *Impatiens* and *Heracleum* but seems unlikely given the long winter period between seed setting and germination for these species. The additive effects of several small responses to climate have been suggested elsewhere to be sufficient to limit species spread (Pigott 1970, Davison 1977). On balance however it seems that such a combination of factors is only likely to restrict *Impatiens* and that the other two species are apparently quite capable of surviving, at least as individuals, beyond their current limit. Population and meta-population effects are more difficult to quantify, indeed more work is yet required to further substantiate effects on individuals.

The climate range in which *Fallopia* exists naturally in Asia has in many respects pre-adapted it for survival in the cool climate of the British Isles, the plants being observed to grow successfully at all sites across the transect. If its inability to set seed at all in Britain is discounted and purely its vegetative spread considered it could be regarded the most successful of these three invasive species. It has been suggested that in Japan survival of current year seedlings determines the species altitudinal limit (Maruta 1983). Its reduced biomass production at higher altitudes may serve to restrict its rate of spread at such sites and its wind intolerance may further limit it to more sheltered areas but other than that it could potentially occur over a much wider areas than it does currently. However the high altitude sites in which *Fallopia* thrives in Asia are characterised by a much warmer growing season (IIASA database, Leemans & Cramer 1991) than the UK upland so it is uncertain whether spread could actually be realised in these areas. *Fallopia* has the potential to become the most pervasive of these alien species and it is probably only its restricted method of spread that currently limits it. Beerling *et al.* (1994) also suggest that human factors are important in its distribution.

A comparison of the heat-sum figures for the sites over the study period with the predictions of occurrence made by Beerling (1993) for *Impatiens* and *Fallopia* indicate that the suitability of sites for the two species varies between years (Figure 7.28). Beerling's limits for *Impatiens* when compared to pod production at the sites seem to tally with a minimum requirement of about 20 pods per plant to permit survival.

The biomass and survival results for *Fallopia* from this work seem to discount the proposal by Beerling (1993) that growing season length is likely to affect the species range due to increased mortality of the smaller rhizome fragments produced at higher altitudes. However it may be that the experimental plants growing from rhizome fragments rely heavily on this store in their first year of growth and in subsequent years may fail if sufficient resources are not acquired. No prediction in terms of accrued biomass necessary for persistence can be reliably suggested for *Fallopia* plants.

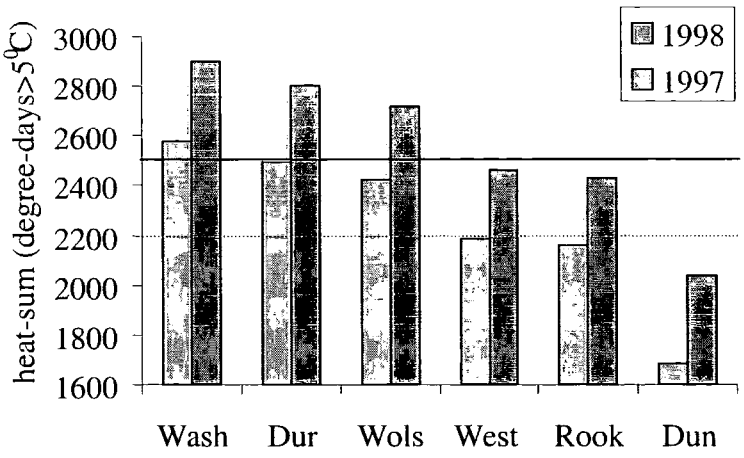

Figure 7.28 Growth season heat-sums of sites (measured as degree-days >5°C) compared to the predicted climatic limits of *Fallopia* (solid line) and *Impatiens* (dashed line).

For all three species the effect of "haar" (coastal fog arising from cool, moist air drawn off the sea) on coastal locations will serve to reduce available sunshine hours (Graham 1988). This may be responsible for the reduced plant productivity at Washington, though in times of summer drought it may have an opposite effect, improving coastal species performance compared to more inland sites.

Although several of the relationships between plant variables and climate are significant and having acknowledged the potential impact of other climate variables there is still a lot of variation between individuals at any site. Some of this variation is likely to result from individual heritable traits but may also be due to factors such as microclimate within the experimental units, slight variations in drainage regime between pots at any site and stochastic herbivory. It must also be remembered that climatic variables are generally correlated with each other (Ford & Milne 1981) and that an apparent relationship between some plant trait and a climatic variable may be due to the variables correlation with other climatic factors.

It should also be remembered that this experiment was designed so as to minimise variation in performance due to factors other than climate. In doing so the plants have been grown in a somewhat artificial environment. For this reason it may not be appropriate to simply scale up these results to create a model for predicting species potential ranges. Upland area for example generally have a very different soil type, pH and drainage regime than do lowland area and such factors need to be taken into consideration before any reliable predictions can be made. The experiments undertaken

here merely indicate where the species could potentially grow given that all other environmental conditions were suitable.

It may be that the occurrence of slug species or rodent species in some habitats may be the influential factor restricting the occurrence of these species in natural situations. Similarly it could be postulated that these species do best in the continually flooded ruderal zones of the riparian environment because of a paucity of predators such as slugs to predate seedlings and rodents to eat the seeds rather than the oft-cited cause of reduced plant species competition.

The loss of the site at Wolsingham in 1997 due to grazing by stray cattle showed that *Impatiens* and *Heracleum* were acceptable as grazing fodder for cattle, *Heracleum* seeming to be particularly favoured. *Fallopia* plants were generally uprooted by the cattle but then left uneaten. This suggests cattle could be used as a tool for the control of *Heracleum* as has been stated elsewhere (Andersen 1994; Dodd *et al.* 1994). Similarly cattle or sheep could be used to control *Impatiens*, both species grazing plants indiscriminately (Beerling and Perrins 1993). These observations back up the analyses carried out using the RCS data, which generally found the two species associated with ungrazed areas and avoiding grazed grassland.

In summary it can be seen that climate has a very marked effect on the performance of all three species but that it is not necessarily restricting their occurrence. As well as climate the importance of dispersal, as well as other factors such as habitat, pollinators and predators/herbivores may also need to be considered in predicting future spread.

Chapter Eight

General Discussion

This chapter inter-relates the major findings of this thesis in determining factors responsible for the distribution of the three riparian weed species. The roles of habitat and climate, the two main areas of study, as limiting factors will be dealt with initially. These two elements will then be assessed to determine their relative importance in the distribution of the study species. The demographic results can then be contrasted with the variables selected in the predictive models to see if the findings from these two disparate approaches accord and if the selected variables concord with sound ecological reasoning. The role of scale in species modelling is then discussed, the discussion concluding by highlighting the applications to which these findings can and have been used along with suggestions for further work.

The work contained in this thesis contains several novel approaches to the study of alien species, by use of contrasting methods of study. The use of the Environment Agency's RCS data provides a good "top-down" approach to the modelling of alien species; riparian plant modelling using such data has not been attempted before. In contrast, the experimental work examining the effects of habitat on plant demography is largely new to the study of these species and provides new "bottom-up" insights into the importance of different habitat types. Similarly the experiments examining climatic gradients, as represented by elevation, have demonstrated novel effects of climate on various plant traits, from germination to plant biomass and seed traits. The modelling approaches largely fulfil the requirement to predict current species distributions but, by combining these with demographic studies, the work also demonstrates some of the mechanisms by which such models may operate, something that other studies often lack (Williamson 1989; Panetta & Mitchell 1991).

8.1 The role of habitat type

The analyses demonstrate that the species consistently occur more often than expected in certain riparian habitats. Both differences in vegetation types and substrate types are shown to be important for *Heracleum* and *Impatiens*. Tolerance of varied substrate types by *Impatiens* and *Heracleum,* and apparent habitat preferences of all three species from the logistic models reflect reported habitat types. All three species show a degree of correlation

in their occurrence which, in combination with similar modelled habitat preferences, suggests they all occupy a similar large-scale riparian niche. It seems from this work that lower, regularly flooded areas of riverbank are particularly susceptible to invasion by these three species. This may be due to low native species diversity (e.g. Tiley *et al.*1996) and hence reduced competition or herbivory. Alternatively they may simply reflect restrictions such as suitable germination niches or deposition zones for seeds/ rhizomes.

Fine scale analyses of the occurrence of the species along the riverbanks indicates that occurrence is not always as expected given dispersal methods. The continual distribution of *Impatiens* in the lower-riparian zone reflects its water-borne dispersal strategy, whereas *Heracleum*'s patchy predominance in this zone is less expected given its floating seeds and high fecundity. The patchy distribution of *Fallopia* is understandable given its vegetative method of spread, as is its predominance in the lower riparian zone.

Abundance patterns counter the suggestion that these species often dominate whole areas of the riverbank to the exclusion of other species (e.g. Roblin 1994), indeed small patches are the norm for both *Fallopia* and *Heracleum*. All three species can however dominate the vegetation locally.

Experiments demonstrate a quite marked effect of habitat type on germination/shoot production for the three species; in all cases ruderal areas produced greater densities of seedlings/shoots than did woodland or grass/herb habitats. Disturbance in the autumn preceding growth was found to have an insignificant effect on the species in all habitats except for *Heracleum* in ruderal areas, where germination was initially greater. Disturbance in relatively undisturbed habitats therefore seems to be less important than anticipated from other sources of information on alien invasions (e.g. Hobbs 1989). Mortality during the growing season was found to vary both between habitats and years. Survivorship was consistently low for *Impatiens* in ruderal areas but more variable in the other two habitats, mean survival time always being greater in woodland than the other habitats. *Heracleum* demonstrated consistently high summer mortality across habitats whereas *Fallopia* had almost 100% survival.

Survivorship of the three species over the entire growing season was found to show only weak density dependence. However when density dependent effects were split into monthly intervals, thereby taking account of plant growth stages, *Impatiens* was found to

demonstrate strong density dependent mortality, especially in ruderal and woodland habitats in the periods May-June and July-August. This was explicable in terms of plant growth stages, e.g. period of leaf expansion.

Seed output of *Impatiens* plants was always greater in ruderal and grass/herb areas than in woodland. Density of plants was also greatest in ruderal areas and lowest in grass/herb vegetation. Similarly *Heracleum* seed set was greatest in ruderal areas, where plants probably occur at increased densities. In ruderal areas both these species demonstrated large seed-set, sufficient to exceed the invasion criteria of Crawley *et al.* (1993).

All of the habitat modelling is to some extent tempered by the poor quality of the RCS data and potentially inaccurate species distributions along the Tees catchment. It may be that with better habitat data and accurate distributions, trends would become more apparent.

8.2 The importance of climate

The early germination dates of plants, particularly *Heracleum* and *Impatiens* reflect a tendency to out-compete other riparian plant species. Reduced numbers of seedlings arising from early germinating *Impatiens* seeds at some sites suggest that small-scale microclimatic events may also be important at this crucial period, as observed elsewhere (Perrins *et al.*1993). Variation in timing of germination of pot grown and wild plants of *Heracleum* and *Impatiens* also suggest that factors other than large-scale climate have an influential role. *Impatiens* was the only species to show a correlation between winter climate and seed germination, and even this relationship was weak, suggesting that the winter climates experienced during the two years of study were not sufficient to result in increased mortality of any of the species' seeds/rhizome fragments. It was concluded that none of the species was limited in their ability to colonise further areas of the catchment due to an inability to germinate from seed/rhizomes. However increased mortality of over-wintering first-year *Heracleum* plants was shown to occur in the more upland areas.

Plant performance of all three species was affected by climate, leading to reduced stature when summer heat sum is low. Two-year old *Heracleum* plants were the only plants to demonstrate retarded growth with excess summer temperatures; an optimum climate of 1700 degree-days over 5°C being suggested. Mass mortality of *Impatiens* plants in several sites during the same period in 1998, irrespective of growth stage, implies a climatic cause, either as a direct effect or an indirect effect on a herbivore/pathogen.

Important effects of climate on *Impatiens* plants were detected. Reduced pod production, later and more limited seed production occurred as summer heat sum declined. Conversely individual seed weights were, on average, greater in areas of impoverished summer climate. With *Heracleum* an examination of seed output from lowland plants compared to those from its upper limit showed no decline in fecundity, suggesting that climate, operating by reducing seed output, is not limiting the spread of this species at increased altitudes.

8.3 What determines distributions?

From the above it can be seen that both habitat and climate are influential in the demography of the study species in County Durham. *Fallopia* and *Heracleum* seem not to be limited by climate, although reduced stature is observed at the most extreme sites. As the habitat in which they occur is also present outside their current range it seems likely that either dispersal or some other factors (e.g. soil types, flood regime, herbivory) are limiting these species. In the case of *Impatiens,* climate is observed to reduce fecundity, which in combination with seed and seedling mortality may be sufficient to produce the current distribution limit. Dispersal must also have a role to play in limiting colonisation of some areas distant from the main river populations. Bankside management could also act as a barrier to dispersal by creating sections of unsuitable habitat, past which species may not easily spread.

Suitable habitat, given the broad RCS categorisations, seems to be available for all three species beyond their current range. It would seem that suitable habitats are so widespread along river systems as to not restrict the species. The reduced performance of plants, in woodland areas in particular, has no great effect on species' persistence in such areas, leading to the conclusion that demographic variation between habitats may not be sufficient to restrict species. Since it has been suggested that individual *Impatiens* plants need produce about 20 pods in order to persist in an area (Chapter 7), then plants in woodland, having pod production figures marginally greater than this value, should be capable of persistence. The effects of long-distance seed dispersal along the rivers may also partially compensate for differing fecundity. The reduced fecundity of the seed producing species along the upper bank habitats may however restrict the spread of the species into the wider environment as a result of low seed output. Similarly reduced disturbance of *Fallopia* stands outside of the flooded zone will restrict spread of rhizome fragments.

The optimal performance of the species in ruderal areas close to the river margin will assist dispersal by facilitating maximum propagule output in areas from which they are easily widely dispersed. As such this habitat plays an important role in dispersal.

Overall therefore it seems that altitude is one of the primary determinants of geographical range for *Impatiens*, whereas for the other two species dispersal may be more important. Habitat type does not seem to be as important in restricting spread but rather ruderal habitats encourage spread generally by making propagules available for long distance transport.

8.4 Comparing models with demographic analyses

Best fit models for the prediction of occurrence of *Fallopia* had relatively few common variables. Of these, variables such as urban areas and roads are probably related to sites of introduction. Others such as solid earth/rock cliffs reflect a habitat the species would be expected to avoid but were not covered in the demographic studies. The selection of woodland rather than ruderal areas as a common variable contradicts demographic findings though it may be that rhizome fragments are more likely to get snagged in such areas or may reflect a reliance on woodland fringe communities.

The logistic models for *Heracleum* consistently select altitude as an important variable, though results from the elevation transect suggest that altitude is not a limiting factor and its selection is determined only by the species currently restricted distribution. As such, predictions derived from such a model, based on a non-equilibrium distribution will prove erroneous in the future. This fact should be borne in mind when modelling alien species spread and has been suggested elsewhere as a potential source of error in producing models (Collingham *et al.* in press; Franklin 1995). Woodland is commonly selected as an influential variable for *Heracleum* whereas ruderal areas are less frequently selected, surprising given the demographic and distribution analyses. The common inclusion of disturbance as a negative factor contradicts the demographic studies herein, which found it not to influence plant numbers, though it must be remembered that disturbance factors may be quite different in terms of duration etc.

The logistic models of *Impatiens* occurrence generally tally with the demographic findings; ruderal areas are selected as preferential habitat. Woodland, whilst important in the several single-catchment models is not so important in the composite models. Altitude is an

important variable, probably correctly so for this species, though the selection of grassland vegetation as a preferred associate conflicts to some extent with the demographic studies. It has also been shown that several habitat variables have differing effects on species' occurrences in distinct zones of the riverbank.

The upper and lower altitudinal limits suggested from the models have been justified by the demographic experiments only for *Impatiens* and the seaward limit of *Heracleum*.

8.5 The role of scale in species modelling

Modelling the distribution of the study species at different scales has provided interesting comparisons. Despite being able to produce good predictive models at the scale of 500m RCS sections within a single catchment (e.g. Willis *et al.* 1997 and Chapter 3) this ability declines when tested on other catchments (Chapter 3). By increasing the scale to that of a 10km grid across Britain, the ability to model successfully species occurrence has been found to decline substantially (Collingham *et al.* 1997). For *Impatiens* at least this has been attributed to the fact that within a catchment it has often realised its potential range, whereas on a national scale this is not necessarily the case (Collingham *et al* 1997). Conversely *Fallopia* models at a catchment scale are very poor, whereas those based on a 50km grid (Beerling *et al.* 1995) produce good fits. It is reasonable to assume that *Fallopia*, having been widely planted, has had chance to naturalize in most 50km squares but had not yet been able to spread to all potential sites at a smaller scale.

8.6 The results in light of published work

The results from the demographic work undertaken as part of this thesis concur with much of the currently available data. There are however several anomalies. The importance of the lower-riparian bankside and the consequences of this on species spread have previously been underplayed. Observed germination densities of *Impatiens* plants are generally higher than given elsewhere, whereas flowering plant densities are low compared to some published data. This highlights the fact that much previous work has been undertaken in areas of optimal growing conditions and that sub-optimal habitats have largely been overlooked, masking actual variability. Reduced performance of *Impatiens* and *Heracleum* in selected habitats has been suggested elsewhere but never quantified. Similarly the effects of altitude on the species performances have not been quantified until now. The differential effects of habitat types in different zones of the riverbank have not been previously noted

and the ability to model abundance of *Impatiens* with some degree of success is novel. Surprisingly an increase in resolution of data collection, from 500m to 50m, does not improve the ability to model species occurrence, suggesting higher ecological resolution may be necessary.

The observed high seedling mortality of *Heracleum* and *Impatiens* tallies with observations elsewhere, though increased over-winter survival of *Heracleum* plants in woodlands suggests that ruderal areas are more prone to wash-out/inundation mortality. Reduced densities of all species in woodland corroborate other work. Mortality rates of *Impatiens* plants concur with the findings of Prach (1994), though the situation was shown to not be so straightforward in all habitats. Reduced mortality of *Impatiens* at lower densities agrees with Perrins *et al.* (1993). The effects of density dependent mortality presented here refine the findings of previous work (Beerling 1990a; Perrins *et al.* 1993) to highlight that density dependent mortality depends upon the growth stage of the plant and does not operate equally throughout the growing season.

The seed production of *Heracleum* and *Impatiens* plants is comparable to other available data (Beerling 1990a; Koenies and Glavec 1979; Neiland 1986; Oschmann 1992).

The altitude experiments demonstrate several novel phenomena. The relationships between germination date and spring frost days and heat sum and also the relationship between biomass/plant height and summer degree-days have not been previously quantified. This work substantiates the logical prediction that occurrence models based on current distributions will not give a true reflection of sites of spread in the near future (e.g. Sindel & Michael 1992).

Results can be viewed in light of work looking at limiting factors for other species. Pollinator limits to spread, as have been observed for other species (Parker 1997; Ingvarsson & Lundberg 1995), cannot be substantiated as a limiting mechanism for the study species. Similarly, as plants have proven capable of setting seed above current limits, fruit development limitations are not apparent (as for example found by Pigott & Huntley 1981). An ability of individuals to produce seed above current limits does not however eliminate the possibility of secondary effects on population or metapopulations. Herbivory as a limiting mechanism also cannot be ruled out; having proven important in limiting other species (e.g. Herrera 1991). Similarly competition effects, which have been shown to be

fundamental in limiting many plant species (e.g. Bockelmann & Neuhaus 1999), cannot yet be quantified for the study species. The various, and often complex (e.g. Woodward 1997), climatic limitations noted for many species, were shown to affect the two seed producing species to varying extents in the study area. Microsite limitation, noted as important for other species (e.g. Hulme 1996), has not been quantified but given the widespread availability of apparently suitable lower-bank habitat this is unlikely to be restricting the species, at least along the river margins.

Differences in pod content of *Impatiens* with differing climate and habitat have not been noted before, neither has the change in seed weight over time and in differing climate. Such findings do however have parallels in work undertaken previously on other plant species (e.g. Cavers & Harper 1966; Fisher & Mallhies 1998). The correlation between seed weight and temperature during the period of pod development is also novel. The potential effects of these findings require further evaluation.

The *Impatiens* pod production values, when compared to Beerling's (1993) calculated climatic limits are found to correspond to a value of about 20 pods per plant being necessary to permit survival. Experiments involving *Fallopia* negate Beerling's (1993) proposal that smaller rhizomes fragments might be expected to have a higher mortality at increased altitude.

Both the observed livestock grazing events and the RCS models support other work, which advocates grazing as a means of control (Andersen 1994; Beerling & Perrins 1993 etc.).

The hierarchy of environmental controls over plant distribution, which has been suggested elsewhere (Franklin 1995), can be viewed in light of the important variables in the various logistic models. It has been suggested that climatic variables determine patterns at the larger scales of study, followed by geology (and its effects on soil chemistry and nutrients) and topography at decreasing scales of study. Although Collingham *et al.* (in press) found no hierarchical effects when comparing tetrad and hectad datasets for these species, the present work, working at smaller scales demonstrates a degree of hierarchical controls. Variables such as altitude and river width operate to initially define a climatic tolerance (partly substantiated by demographic studies) with habitat variables subsequently incorporated into the models to refine the predicted distributions.

One of the most important conclusions of the work, namely that two of the three species are probably not climatically limited (at least within the study area) is of importance with regards predictive modelling. Other work produced by the group (Wadsworth *et al.*, submitted) has suggested that along the Wear catchment there is an optimum time for seed dispersal: namely the period between late August and mid-October, though the differences in median and maximum dispersal distances are slight. For both *Impatiens* and *Heracleum* almost twice the number of seeds would "escape" from a section during this period compared to earlier in the year. It may therefore be the case that the earlier production of *Impatiens* seeds in the lower areas of the catchment results primarily in the re-establishment of plants within a currently colonised section. The seeds from the upper areas of catchment, being dispersed later, may theoretically be responsible for much of the long-distance dispersal. If this is so (and this assumes that early seeds once deposited are immobile) then seed introduced into the upper areas of a catchment will result in a more rapid colonisation of a downstream area than introductions in lower parts of the catchment. Similarly control of upper catchment populations would be most effective in reducing population spread.

Predicted seed dispersal distances for *Impatiens* and *Heracleum* produced by Wadsworth *et al.* (submitted) highlights the fact that *Heracleum* with a lower seed dispersal distance would be the easier of the two to control, recolonisation of distant sites being less likely.

The apparent rapid recent consolidation and spread of the species, notably *Fallopia*, on the Wear substantiates advocating a rapid response to deal with introduced aliens, thereby eliminating problems before they become a) too costly and b) too widespread to deal with.

Modelling work by Wadworth *et al.* (in press), utilising demographic and habitat parameters produced as part of the work of this thesis, has highlighted the importance of adopting a reliable control strategy in order to eradicate the study species. As well as advocating a control strategy based on the size of plant populations or starting at the head of a catchment and working downstream, the work emphasizes the importance of a sufficient intensity of control, in terms of area of control and also efficiency of control within areas.

8.7 Further work

It has been shown that climate plays an important role in determining species' distributions though, because of correlation between climatic variables (Ford & Milne 1981), the exact

mechanisms of limitation can only be speculated upon. The use of controlled climate experiments to examine growth parameters under differing conditions would be of use in predictive modelling, especially if individual variables could be manipulated. Also, as climate has been shown to alter plant growth and reproductive traits, it would be useful to investigate how these altered traits affect species survival over longer periods. Would, for example, the reduced biomass of *Heracleum* plants at altitude result in an increased time to flowering or reduced seed output? The variation in seed weight of *Impatiens* plants with altitude would also benefit further work. Do large seeds survive better at altitude than smaller seeds, and what are the effects of differing growth mediums on seed survival?

It would be useful to be able to sow seeds/rhizome fragments of the species into the natural environment to examine survival more accurately in riparian habitats and substrates. The work undertaken as part of the habitat analyses in this thesis is inherently limited in its inability to assess seed inputs and seed mortality. Work on seed banks, both in terms of survival times and also in dispersal patterns would provide data useful for both modelling and control programmes and would fill a grey area in the ecology of the species.

Although the work presented here shows that habitat types have an important effect on the demography of alien species and that different habitats are preferentially selected, the habitats themselves are very broadly defined. Further investigations examining species demography in relation to more detailed habitat types would be worthwhile. Such differing effects are masked in the broad-brush approach, necessarily adopted here. The work using habitat traits to produce predictive models would also benefit from constructing models using a more diverse and widely distributed set of catchments. The work of Collingham *et al.* (submitted) has already shown that models based on a wider distribution are better predictors at a small scale than regional models are at predicting on a national scale.

References

ACRE (1999). "Department of the Environment, transport and the regions." Advisory Committee on Release to the Environment (ACRE) *hhtp://www.environment.detr.gov.uk/acre/wildlife/02.htm.*

Allen, L.J.S., Allen, E.S., Kunst, C.R.G. & Sosebee, R.E. (1991). "A diffusion model for dispersal of *Opuntia imbricata* (Cholla) on rangeland." *Journal of Ecology* **79**: 1123-1135.

Andersen, U.V. (1994). "Sheep grazing as a method of controlling *Heracleum mantegazzianum*". In: *Ecology and management of invasive riverside plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 77-91.

Andrews, A.H., Giles, C.J. & Thomsett, L.R. (1985). "Suspected poisoning of a goat by giant hogweed." *Veterinary Record* **116**: 205-207.

Angold, P.G., Gurnell, A.M. & Edwards, P.J. (1996). "Locational errors in maps from environmental surveys and their implications for information extraction." *Journal of Environmental Management* **47**: 341-354.

Ashton, P.J. & Mitchell, D.S. (1989). Aquatic plants: patterns and modes of invasion, attributes of invading species and assessment of control programmes. In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons. 111-154.

Atkinson, I.A.E. (1985). The spread of *Rattus* to oceanic islands and their effects on island avifaunas. In: *Conservation of island birds.* **Ed**.Moors, P.J. International Council for Bird Preservation. **Technical Publication 3:** 35-81.

Augustin, N.H., Mugglestone, M.A. & Buckland, S.T. (1996). "An autologistic model for the spatial distribution of wildlife." *Journal of Applied Ecology* **33**: 677-692.

Bailey, J.P. (1988). "Putative *Reynoutria japonica* Houtt. x *Fallopia baldshuanica* (Regel) Holub. hybrids discovered in Britain." *Watsonia* **17**: 163-181.

Bailey, J.P. & Conolly, A.P. (1991). "Alien species of *Polygonum* and *Reynoturia* in Cornwall 1989-1990." *Botanical Cornwall Newsletter* **5**: 33-46.

Baker, H.G. (1965). "Characteristics and modes of origin of weeds". In: *The genetics of colonising species*. **Eds**: Baker, H.G. & Stebbins, G.L.. London. Academic Press. 147-168.

Baker, H.G. (1986). "Patterns of plant invasion in North America". In: *Ecology of biological invasions of North America and Hawaii*. **Eds**: Mooney, H.A. & Drake, J.A.. New York. Springer-Verlag. 44-57.

Barrow, E., Hulme, M. & Jiang, T. (1993). *A 1961-90 baseline climatology and future climatic change scenario for Great Britain and Europe. Part 1: 1961-90 Great Britain baseline climatology*. Norwich. Climate Research Unit, University of East Anglia.

Beerling, D.J. (1990a). *The ecology and control of Japanese Knotweed and Himalayan Balsam on river banks in South Wales*. Ph.D. Thesis. Cardiff, University of Wales.

Beerling, D.J. (1990b). "The use of non-peristent herbicides, glyphosphate, and 2,4-D amine, to control riparian stands of Japanese Knotweed (*Reynoutria japonica* Houtt.)." *Regulated rivers: Research and management* **5**: 413-417.

Beerling, D.J. (1991a). "The effect of riparian land-use on the occurrence and abundance of Japanese Knotweed, *Reynoutria japonica*, on selected rivers in South Wales." *Biological Conservation* **55**: 329-337.

Beerling, D.J. (1991b). "The testing of cellular revetment blocks resistant to the growth of *Reynoutria japonica* Houtt. (Japanese Knotweed)." *Water Resources* **25**(4): 495-498.

Beerling, D.J. (1993). "The impact of temperature on the northern distribution limits of the introduced species *Fallopia japonica* and *Impatiens glandulifera* in north-west Europe." *Journal of Biogeography* **20**: 45-53.

Beerling, D.J. (1994). Predicting the response of the introduced species *Fallopia japonica* and *Impatiens glandulifera* to global climatic change. In *Ecology and Management of Invasive Riverside Plants*. **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 135-139.

Beerling, D.J., Bailey, J.P. & Conolly, A.P. (1994). "Biological Flora of the British Isles: *Fallopia japonica* (Houtt.) Ronse Decraene (*Reynoutria japonica* Houtt.; *Polygonium cuspidatum* Sieb & Zucc.)." *Journal of Ecology* **82**: 959-979.

Beerling, D.J., Huntley, B. & Bailey, J.P. (1995). "Climate and the distribution of *Fallopia japonica*: use of an introduced species to test the predictive capacity of response surfaces." *Journal of Vegetation Science* **6**: 269-282.

Beerling, D.J. & Palmer, J.P. (1994). Status of *Fallopia japonica* (Japanese Knotweed) in Wales. In *Ecology and Management of Invasive Riverside Plants*. **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 199-211.

Beerling, D.J. & Perrins, J.M. (1993). "Biological Flora of the British Isles: *Impatiens glandulifera* Royle (*Impatiens roylei* Walp.)." *The Journal of Ecology* **81**: 367-382.

Beerling, D.J. & Woodward, F.I. (1994). "Climate change and the British scene." *Journal of Ecology* **82**: 391-397.

Birks, H.J.B. (1973). *Past and present vegetation of the Isle of Skye. A paleoecological study*. Cambridge. Cambridge University Press.

Bockelmann, A.C. & Neuhaus, R. (1999) Competitive exclusion of *Elymus athericus* from a high-stress habitat in a European salt marsh. *Journal of Ecology* **87** (3): 503-513.

Boudouresque, C.F., Briand, F. & Nolan, C. (1994). *Introduced species of European coastal waters: international workshop report*. The Commission.

Bradshaw, M.E. (1970). "The Teesdale flora". In *Durham County and City with Teeside*. **Ed**. Dewdney, J.C. Durham. 141-152.

Braun-Blankuet, J. (1932). *Plant sociology: The study of plant communities (English translation)*. New York. McGraw-Hill.

von Broembsen, S.L. (1989) "Invasions of natural ecosystems by plant pathogens". In: *Biological invasions: a global perspective*. **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons. 77-84.

Brondegaard, V.J. (1990). "Massenausbreitung des barenklaus." *Naturwissenschaftliche Rundschau* **43** (10): 438-439.

Brown, J.H. (1989). "Patterns, modes and extents of invasions by vertebrates". In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M.. Chichester. John Wiley and Sons. 85-109.

Buckland, S.T. & Elston, D.A. (1993). "Empirical models for the spatial distribution of wildlife." *Journal of Applied Ecology* **30**: 478-495.

Buckland, S.T., Elston, D.A. & Beaney, S.J. (1996). "Predicting distributional change, with application to bird distributions in northeast Scotland." *Global Ecology and Biogeography Letters* **5**: 66-84.

Buckton, S.T. & Ormerod, S.J. (1997). "Use of a new standardised habitat survey for assessing the habitat preferences and distribution of upland river birds." *Bird Study* **44**: 327-337.

Butler, D. & Reichhard, T. (1999). "Long-term effect of GM crops serves up food for thought." *Nature* **398** (22 Apr): 651-656.

Caffrey, J.M. (1994). "Spread and mangement of *Heracleum mantegazzianum* (Giant Hogweed) along Irish river corridors". In: *Ecology and management of invasive riverside plants.* **Eds**:Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 67-76.

Camm, E., Buck, H.W.L. & Mitchell, J.C. (1976). "Phytophotodermatitis from *Heracleum mantegazzianum.*" *Contact Dermatitis* **2**: 68-72.

Catchpole, A.J.W. (1963). "The Houghall frost hollow." *Met. Magazine.* **92**: 121-129.

Cavers, P.B. & Harper, J.L. (1966). "Germination polymorphism in *Rumex crispus* and *Rumex obtusifolius.*" *Journal of Ecology* **54**: 367-82.

CHEST (1993). *Bartholomew Digital Map Data 1993 GB 1:250,000.* University of Bath.

Child, L. & de Waal, L. (1997) "The use of GIS in the management of *Fallopia japonica* in the urban environment." In *Plant invasions: studies from North America and Europe.* **Eds.** Brock, J.H., Wade, M, Pysek, P. & Green, D. Bachhuys Publishers, Leiden. 207-220.

Clegg, L.M. & Grace, J. (1974). "The distribution of *Heracleum mantegazzianum* (Somm. and Levier) near Edinburgh." *Transactions of the Botanical Society of Edinburgh* **42**: 223-229.

Clements, E.J. & Foster, M.C. (1994). *Alien plants of the British Isles.* London. Botanical Society of the British Isles.

Cliff, A.D. & Ord, J.K. (1981). *Spatial processes: models and applications.* London. Pion Limited.

Coblentz, B.E. (1978). "The effects of feral goats (*Capra hircus*) on island ecosystems." *Biological Conservation* **13**: 279-286.

Colasanti, R.L. & Grime, J.P. (1993). "Resource dynamics and vegetation processes: a deterministic model using two-dimensional cellular automata." *Functional Ecology* **7**: 169-176.

Collingham, Y.C., Huntley, B. & Hulme, P.E. (1997). "A spatially-explicit model to simulate the spread of a riparian weed". In: *Species dispersal and land-use processes: Proceedings of the sixth annual IALE (UK) conference.* Colraine, University of Ulster. IALE (UK).

Collingham, Y.C., Wadsworth, R.A., Huntley, B. & Hulme, P.E. (in press). "Predicting the spatial distribution of alien riparian species: issues of spatial scale." *Journal of Applied Ecology.*

Connell, J.H. & Slatyer, R.O. (1977). "Mechanisms of succession in natural communities and their role in community stability and organisation." *American Naturalist* **111**: 1119-1144.

Conolly, A.P. (1977). "The distribution and history in the British Isles of some alien species of *Polygonum* and *Reynoutria.*" *Watsonia* **11**: 291-311.

Cook (1985). "Range extensions of aquatic vascular plant species." *Journal of Aquatic Plant Management* **23**: 1-6.

Coombe, D.E. (1956). "Biological flora of the British Isles: *Impatiens parviflora* DC." *Journal of Ecology* **44**: 701-713.

Cowie, I.D. & Werner, P.A. (1993). "Alien plant species invasive in Kakadu National Park, tropical northern Australia." *Biological Conservation* **63**: 127,135.

Crawley, M.J. (1987). "What makes a community invasible?" In: *Colonisation, succession, stability and diversity.* **Eds**: Gray, A.J., Crawley, M.J. & Edwards, P.J. Oxford. Blackwell Scientific Publications. 429-453.

Crawley, M.J. (1989a). "Chance and timing in biological invasions". In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons. 407-424.

Crawley, M.J. (1989b). "Invaders." *Plants Today* (September-October): 152-158.

Crawley, M.J., Hails, R.S., Rees, M., Kohn, D. & Buxton, J. (1993). "Ecology of transgenic oilseed rape in natural habitats". *Nature* **363**: 620-623

Darwin, C. (1872). *The origin of species.* London. John Murray.

Davidson, A.W. (1977) The ecology of *Hordeum murinum* L. III. Some effects of adverse climate. *Journal of Ecology* **58**: 453-466

DeBach, P. (1974). *Biological control by natural enemies.* London. Cambridge University Press.

DeCamps, H. & Tabacchi, E. (1994). "Species richness in vegetation along river margins." In: *Aquatic ecology: scale, pattern and process.* **Eds**: Giller, P.S., Hildrew, A.G. & Raffaelli, D.G.. Cambridge. Blackwell Science. 1-20.

Devine, R.S. (1998). *Alien invasion: America's battle with non-native animals and plants.* U.S.A. National Geographic.

di Castri, F. (1989). "History of biological invasions with special emphasis on the Old World". In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons. 1-30.

di Castri, F., Hansen, A.J. & Debussche, M. (1990). *Biological invasions in Europe and the Mediterranean Basin.* London. Kluwer Academic.

Dixon, M.D. & Johnson, W.C. (1999). "Riparian vegetation along the middle Snake River, Idaho: zonation, geographical trends and historical changes." *Great Basin Naturalist* **59**(1): 18-34.

Dodd, F.S., de Waal, L.C., Wade, M.P. & Tiley, G.E.D. (1994). "Control and management of *Heracleum mantegazzianum* (Giant Hogweed)". In: *Ecology and Management of Invasive Riverside Plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 111-126.

Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M., Eds. (1989). *Biological invasions: a global perspective.* Chichester, John Wiley and Sons.

Drever, J.C. & Hunter, J.A.A. (1970). "Giant hogweed dermatitis." *Scottish Medical Journal* **15**: 315-319.

Dunham, K.C. (1948). *Geology of the northern Pennines ore-field. I. Tyne to Stainmore.* London. Mem. Geol. Surv. HMSO.

Dunn, T.C. (1977). "Notes: pollination of Himalayan balsam." *The Vasculum* **62**(4): 61.

Dynesius, M. & Nilsson, C. (1994). "Fragmentation and flow regulation of river systems in the northern third of the world." *Science* **266**: 753-762.

Ellenberg, H. (1988). *Vegetation ecology of Central Europe.* Cambridge. Cambridge University Press.

Elmore, C.D. & Paul, R.N. (1983). "Composite list of C4 weeds." *Weed Science.* **31**: 686-692.

Elton, C.S. (1958). *The ecology of invasions by animals and plants*. London. Metheun.

Eriksson, O. (1999). "Seed size variation and its effect on germination and seedling performance in the clonal herb, *Convallaria majalis*." *Acta Oecologica-International Journal of Ecology* **20** (1): 61-66.

Fischer, M. & Mallhies, D. (1998). "Effects of population size on performance in the rare plant *Gentianella germanica*." *Journal of Ecology* **86** (2): 195-204.

Ford, E.D. & Milne, R. (1981). "Assessing plant responses to the weather". In: *Plants and the atmospheric environment*. **Eds**: Grace, J., Ford, E.D. & Jarvis, P.G. British Ecological Society. 333-362.

Franklin, J. (1995) "Predictive vegetation mapping: geographical modelling of biospatial patterns in relation to environmental gradients." *Progress in Physical Geography* **19**: 474-499.

Gibbs, J.N., Lipscombe, M.A. & Pearce, A.J. (1999). "The impact of *Phytophthora* disease on riparian populations of common alder (*Alnus glutinosa*) in southern Britain." *European Journal of Forest Pathology* **29**(1): 39-50.

Gibson, S.W., Tiley, G.E.D. & Philp, B. (1995). *The impact of invasive weeds on riverine flora in Scotland: a preliminary report*. Edinburgh. Scottish Natural Heritage.

Grace, J. (1981). "Some effects of wind on plants". In: *Plants and the atmospheric environment*. **Eds**: Grace, J., Ford, E.D. & Jarvis, P.G. British Ecological Society. 31-56.

Grace, J. & Nelson, M. (1981) "Insects and their pollen loads at a hybrid *Heracleum* site." *New Phytologist* **87**: 413-423

Graham, G.G. (1988). *The flora and vegetation of County Durham*. Durham County Conservation Trust.

Greene, D.F. & Johnson, E.A. (1995). "Long-distance wind dispersal of tree seeds." *Canadian Journal of Botany* **73**: 1036-1045.

Green, R.E., Osborne, P.E. & Sears, E.J. (1994). "The distribution of passerine birds in hedgerows during the breeding season in relation to characteristics of the hedgerow and adjacent farmland." *Journal of Applied Ecology* **31**: 677-692.

Gregory, R.D., Carter, S.P. & Baillie, S.R. (1997). "Abundance, distribution and habitat use of breeding goosanders *Mergus merganser* and red-breasted mergansers *Mergus serrator* on British rivers." *Bird Study* **44**: 1-12.

Greig-Smith, P. (1957). *Quantitative plant ecology*. London. Butterworth.

Greig-Smith, P. (1961). "The use of pattern analysis in ecological investigations." *Recent Advances in Botany* **2**: 1354-1358.

Grime, J.P., Hodgson, J.G. & Hunt, R. (1988). *Comparative plant ecology : A functional approach to common British species*. London. Unwin Hyman.

Grime, J.P., Mason, G., Curtis, A.V., Rodman, J.B., Mowforth, M.A.G., Neal, A.M. & Shaw, S. (1981). "A comparative study of germination characteristics in a local flora." *Journal of Ecology* **69**: 1017-1059.

Groves, R.H. & Burdon, J.J., Eds. (1986). *Ecology of biological invasions: an Australian perspective*. Canberra, Australian Academy of Science.

Groves, R.H. & di Castri, F., Eds. (1991). *Biogeography of Mediterranean invasions*. Cambridge, Cambridge University Press.

Gunn, I.D.M. (1986). *A distributional survey of Japanese knotweed (Reynoutria japonica Houtt.) and Himalayan balsam (Impatiens glandulifera Royle)*. M.Sc. Thesis, University of Wales, Cardiff.

Gupta, R.K. (1989) *The Living Himalayas. Volume 2: Aspects of plant exploration and phytogeography*. New Delhi. Today and Tomorrows Printers and Publishers. (quoted from **Beerling, D.J. & Perrins, J.M.**, 1993)

Gurnell, A.M. (1995). "Vegetation along river corridors: hydrogeomorphological interactions." In: *Changing river channels*. **Eds**: Gurnell, A.M. & Petts, G. Chichester. John Wiley and sons. 237-260.

Gurnell, A.M. & Gregory, K.J. (1987) "Vegetation characteristics and the prediction of run-off: analysis of an experiment in the New Forest, Hampshire. *Hydrological Processes* **1**: 125-142

Hanson, J.S., Malanson, G.P. & Armstong, M.P. (1990). "Landscape fragmentation and dispersal in a model of riparian forest dynamics." *Ecological Modelling* **49**: 277-296.

Harper, D., Witkowski, F., Kemp-McCarthy, D. & Crabb, J. (1997). "The distribution and abundance of riparian trees in English lowland floodplains." *Global Ecology and Biogeography Letters* **6**: 297-306.

Harper, J.L. (1977). *Population biology of plants*. London. Academic Press.

Harper, J.L., Lovell, P.H. & Moore, K.G. (1970). "The shapes and sizes of seeds." *Annals of Ecology and Systematics* **1**: 327-356.

Harrison, S. (1994). "Metapopulations and conservation". In: *Large-scale ecology and conservation biology*. **Eds**: Edwards, P.J., May, R.M. & Webb, N.R.. Oxford. Blackwell Science. **35.** 111-128.

Harwood, D.G. (1985). "Giant hogweed and ducklings." *Veterinary Record* **March 16th**: 100.

Hegi, G. (1906-31). *Illustrierte Flora von Mittel-Europa.* Ed 1.Vol. 5 Munchen, Germany

Hengeveld, R. (1994). "Small-step invasion research." *Trends in Ecology and Evolution* **9**(9): 339-342.

Herrera, J. (1991) "Herbivory, seed dispersal and the distribution of a ruderal plant living in a natural habitat". *Oikos* **62** (2): 209-216

Heywood, V.H. (1989). "Patterns, extents and modes of invasions by terrestrial plants". In: *Biological invasions: a global perspective*. **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons.

Hill, N.M., Keddy, P.A. & Wisheu, I.C. (1998). "A hydrological model for predicting the effects of dams on the shoreline vegetation of lakes and reservoirs." *Environmental Management* **22**(5): 723-736.

Hinch, S.G., Somers, K.M. & Collins, N.C. (1994). "Spatial autocorrelation and assessment of habitat-abundance relationships in littoral zone fish." *Canadian Journal of Fisheries Aquatic Science* **51**: 701-712.

Hobbs, R.J. (1989). "The nature and effects of disturbance relative to invasions". In: *Biological invasions: a global perspective*. **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons.

Holm, L.G., Plucknett, D.L., Pancho, J.V. & Herberger, J.B. (1977). *The worst weeds: distribution and biology*. Honalulu. University Press of Hawaii.

Holmes, N., Boon, P. & Rowell, T. (1999). *Vegetation communities of British rivers: a revised classification*. Peterborough. Joint Nature Conservation Committee.

Holmes, N.T.H. & Whitton, B.A. (1977a). "Macrophytes of the River Wear 1966-1976." *The Naturalist*: 53-73.

Holmes, N.T.H. & Whitton, B.A. (1977b). "The macrophytic vegetation of the River Tees in 1975: observed and predicted changes." *Freshwater Biology* **7**: 43-60.

Hugh, D.F. & Holland, D. (1998). *Developments in the distribution and control of some invasive plants in bankside habitats*. 10th EWRS symposium on aquatic weeds. Lisbon.

Hulme, P.E. (1996) "Natural regeneration of yew (*Taxus baccata* L): Microsite, seed or herbivore limitation?" *Journal of Ecology*, **84** (6): 853-861

Huntley, B., Berry, P.M., Cramer, W. & McDonald, A.P. (1995). "Modelling present and potential future ranges of some European higher plants using climate response surfaces." *Journal of Biogeography* **22**: 967-1001.

Huntley, B., Baxter, R., Lewthwaite, K.J., Willis, S.G. & Adamson, J.K. (1998). "Vegetation responses to local climatic changes induced by a water storage reservoir." *Global Ecology and Biogeography Letters* **7**: 241-257.

Ingvarsson, P.K. & Lundberg, S. (1995) "Pollinator functional-response and plant-population dynamics - pollinators as a limiting resource". *Evolutionary Ecology* **9** (4): 421-428

Janssen, K. (1994). *Alien species in the marine environment: introduction to the Baltic Sea and Swedish west coast.* Solna. Swedish Environmental Protection Agency.

Jarvinen, A. (1992). "Spatial patterns of nest-box occupancy in the pied flycatcher *Ficedula hypoleuca* in a mountain birch forest." *Ornis Fennica* **69**: 13-18.

Jensen, J. (1976). "Conservation problems in eutrophic fens". In: *Drift og Pleja af Vade Omrader i de Nordiske Lande.* **Eds**: Jensen, A. & Oversen, C.H. University of Aarhus. **Reports from the Botanical Institute No. 3**.

Jermy, A.C. & Crabbe, J.A. (1978). *The island of Mull: a survey of its flora and environment.* London. British Museum (Natural History).

JNCC (1993). *Handbook for phase 1 habitat surveys - a technique for environmental audit.* Peterborough. Nature Conservancy Council.

Johnston, G.A.L. *et al.* (1970). "Geology of Durham County." *Transactions of the Natural History Society of Northumbria* **41**(1): 5-152.

Johnstone, I.M. (1986). "Plant invasion windows: a time based classification of invasion potential." *Biological Review.* **61**: 369-394.

Kershaw, K.A. (1957). "The use of cover and frequency in the detection of plant communities." *Ecology* **38**: 291-299.

Kershaw, K.A. (1960). "The detection of pattern and association." *Journal of Ecology* **48**: 233-242.

King, L.J. (1966). *Weeds of the world: biology and control.* New York. Interscience Publishers.

Koenies, H. & Glavec, V. (1979). "Uber die Konkurrenzfahigkeit des indischen Springkrautes (*Impatiens glandulifera* Royle) an Fuldaufer bei Kassel." *Philippa* **44**: 47-59.

Kornas, J. (1983). "Man's impact on the flora and vegetation in Central Europe". In: *Man's impact on vegetation.* **Eds**: Holzner, W., Werger, M.J.A. & Ikusima, I. The Hague. Dr. W. Junk Publ. 277-286.

Kruse, C.G., Hubert, W.A. & Rahel, F.J. (1997). "Geomorphic influences on the distribution of yellowstone cutthroat trout in the Absaroka Mountains, Wyoming." *Transactions of the American Fisheries Society* **126**: 418-427.

Lattin, J.D. & Oman, P. (1983). Where are the exotic insect threats? In *Exotic plant pests and North American agriculture.* **Eds**: Graham, C. & Wilson, C. New York. Academic Press: 93-137.

Leemans, R. & Cramer, W. (1991). *The IIASA database for mean monthly values of temperature, precipitation and cloudiness of a global terrestrial grid.* Luxenburg. IIASA.

Legendre, P. (1993). "Spatial autocorrelation: trouble or new paradigm?" *Ecology* **74**(6): 1659-1673.

Legendre, P. & Fortin, M. (1989). "Spatial pattern and ecological analysis." *Vegetatio* **80**: 107-138.

Lever, C. (1994). *Naturalised animals: the ecology of successfully introduced species.* London. Poyser.

Levins, R. (1970). "Extinction." *Lectures on Mathematics in the Life Sciences* **2**: 75-107.

Lohmeyer, W. (1970). "Uber das Polygono-Chenopodietum in Westdeutchland unter besonderer Berucksichtigung seiner Vorkommen am Rhine und im Mundungsgbiet der Ahr." *Schriften der Vegetationskunde* **5**: 7-28.

Loope, L.L. (1992). "An overview of problems with introduced plant species in national parks and biosphere reserves of the United States". In: *Alien plant invasions in native ecosystems of Hawaii: management and research.* **Eds**: Stone, C.P., Smith, C.W. & Tunison, J.T. Honalulu. Hawaii Univ. Press.

Lundström, H. (1984). *Giant Hogweed, Heracleum mantegazzianum, a threat to the Swedish countryside.* Weeds and Weed Control, 25th Swedish Weed Conference. **1**, 191-200. Uppsala.

Lundström, H. (1989). *New experiences of the fight against the Giant Hogweed Heracleum mantegazzianum.* Weeds and Weed Control, 30th Swedish Weed Conference. **2**: 51-58. Uppsala.

Macdonald, I.A.W., Loope, L.L., Usher, M.B. & Hamann, O. (1989). "Wildlife conservation and the invasion of nature reserves by introduced species: a global perspective". In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons.

Macdonald, I.A.W., Kruger, F.J. & Ferrar, A.D., Eds. (1986). *The ecology and management of biological invasions in Southern Africa.* Cape Town, Oxford University Press.

Mandenova, I.P. (1950), in Tiley *et al.* 1996. "Caucasian species of the genus *Heracleum.*" *Monographs Series A, Systematics and Plant Geography (in Russian).*

Manley, G. (1936). "The climate of the northern Pennines, the coldest part of England." *Q.J.R.Meteorol.Soc.* **62**: 103-115.

Manley, G. (1941). "The Durham meteorological record 1847-1940." *Q.J.R. Meteorol. Soc.* **67**: 363-380.

Manley, G. (1942). "Meteorological observations on Dunn Fell." *Q.J.R. Meteorol. Soc.* **68**: 151-165.

Manley, G. (1943). "Further climatalogical averages for the northern Pennines, with a note on topographic effects." *Q.J.R. Meteorol. Soc.* **103**: 241-263.

Marchant, J.H. (1983). *BTO common bird census instructions.* Tring. British Trust for Ornithology.

Marsh, G.P. (1865). *Man and nature; or physical geography as modified by human action.* New York. Scribner.

Marshall, J.K. (1968) Factors limiting the survival of *Corynephorus canescens* (L.) Beauv. in Great Britain at the northern limit of its distribution. *Oikos* **19**: 206-216.

Maruta, E. (1983). "Growth and survival of current year seedlings of *Polygonum cuspidatum* at the upper distributional limit on Mt. Fuji." *Oecologia* **60**: 316-320.

Merry, D.J., Slater, F.M. & Randerson, P.F. (1981). "The riparian vegetation of the River Wye." *Journal of Biogeography* **8**: 313-327.

Mitchell, D.S., Ed. (1974). *Aquatic vegetation and its use and control.* Paris, UNESCO.

Mooney, H.A. & Drake, J.A., Eds. (1986). *Ecology of biological invasions of North America and Hawaii.* Ecological Studies. New York, Springer-Verlag.

Moor, M. (1958). "Plantzengesellschaften schweizerischer Flussauen." *Mitt Schweiz Anst. Forstl. Versuchswes* **34**: 221-360.

Mortimer, A.M. (1989). "The biology of weeds". In: *Weeds control handbook: principles.* **Eds**: Hance, R.J. & Holly, K.. Oxford. Blackwell Scientific. 1-42.

Morton, J.K. (1978). "Distribution of the giant cow parsnip (*Heracleum mantegazzianum*) in Canada." *Canadian Field Naturalist* **92** (2): 182-185.

Morton, A.J. (1998). *DMAP: Software for distribution mapping and coincidence mapping.* A.J.Morton, Chawridge Lane, Winkfield, Windsor.

Myers, K. (1986). "Introduced vertebrates in Australia, with emphasis on the mammals". In: *Ecology of biological invasions: an Australian perspective.* **Eds**: Groves, R.H. & Burdon, J.J.. Canberra. Autsralian Academy of Science. 120-136.

Nadeau, S., Decaire, R., Lambert, D. & St-Georges, M. (1995). "Nonlinear modelling of muskrat use of habitat." *Journal of Wildlife Management* **59** (1): 110-117.

Narumalani, S., Jensen, J.R., Althausen, J.D., Burkhalter, S. & Mackey, H.E. Jr. (1997). "Aquatic macrophyte modelling using GIS and logistic multiple regression." *Photogrammatic Engineering and Remote Sensing* **63** (1): 41-49.

National Research Council (1996). *Stemming the tide: controlling introductions of non-indigenous species by ships' ballast water.* U.S.A. National Academic Press.

N.R.A. (1992). *River corridor surveys: Methods and procedures.* Bristol. National Rivers Authority.

Neiland, M.R.M. (1986). *The distribution and ecology of Giant Hogweed (Heracleum mantegazzianum) on the River Allen and its control in Scotland.* Ph.D. thesis. Stirling. University of Stirling.

Neiland, R., Proctor, J. & Secton, R. (1987). "Giant Hogweed (*Heracleum mantegazzium* Somm. & Lev.) by the River Allen and part of the River Forth." *Forth Naturalist and Historian* **9**: 51-56.

Newsome, A.E. & Noble, I.R. (1986). "Ecological and physiological characteristics of invading species." In: *Ecology of biological invasions: an Australian perspective.* **Eds**: Groves, R.H. & Burdon, J.J. Canberra. Australian Academy of Science. 1-20.

Nilsson, C., Gardfjell, M. & Grelsson, G. (1991). "Importance of hydrochory in structuring plant communities along rivers." *Canadian Journal of Botany* **69**(12): 2631-2633.

Nilsson, C. & Grelsson, G. (1990). "The effects of litter displacement on riverbank vegetation." *Canadian Journal of Botany* **68**(4): 735-741.

Noble, I.R. & Slatyer, R.O. (1980). "The use of vital attributes to predict successional change in plant communities subject to recurrent disturbances." *Vegetatio* **43**: 5-21.

Norusis, M.J. (1994). *SPSS Advanced Statistics 6.1.* Chicago. SPSS Inc.

Oberdorfer, E. (1979). *Pflanzensoziologische exkursions flora.* Stuttgard. Verlang Eugen Ulmer.

Ochsmann, J. (1992), in Tiley *et al.* (1996). *Riesen-barenklau (Heracleum sp.) in Deutchland: morphologie, omkologie, verbreitung unt systematische einordnung.* Gotingen. Gotingen University.

Palmer, J.M. (1994). "*Fallopia japonica* (Japanese Knotweed) in Wales". In: *Ecology and management of invasive riverside plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 159-171.

Panetta, F.D. & Mitchell, N.D. (1991). "Bioclimatic prediction of the potential distribution of some weed species prohibited entry to New Zealand." *New Zeal. J. Agr. Res.* **34**: 341-350.

Parker, D.E., Legg, T.P. & Folland, C.K. (1992). "A new daily central England temperature series, 1772-1991." *International Journal of Climatology* **12**: 317-342.

Parker, I.M. (1997) "Pollinator limitation of *Cytisus scoparius* (Scotch Broom), an invasive exotic shrub." *Ecology* **78** (5): 1457-1470

Payne, A.I. (1986). *The ecology of tropical lakes and rivers.* Chichester. John Wiley and Sons.

Perring, F.H., Sell, P.D., Walters, S.M. & Whitehouse, H.L.K. (1964). *A flora of Cambridgeshire.* Cambridge. Cambridge University Press.

Perrins, J., Fitter, A. & Williamson, M. (1993). "Population biology and rates of invasion of three introduced *Impatiens* species in the British Isles." *Journal of Biogeography* **20**: 33-44.

Pigott, C.D. (1956). "The vegetation of Upper Teesdale in the North Pennines." *Journal of Ecology* **44**: 545-586.

Pigott, C.D. (1970) The responses of plants to climate and climate change. *The flora of a changing Britain.* Ed. **Perring, F.H.** 32-34. Botanical Society of the British Isles, Classey. Faringdon.

Pigott, C.D. & Huntley, J.P. (1981). "Factors controlling the distribution of *Tilia cordata* at the northern limits of its geographical range. III Nature and cause of seed sterility." *New Phytologist* **87**: 817-839.

Pigott, C.D. & Wilson, J.F. (1978). "The vegetation of North Fen at Esthwaite in 1967-9." *Proceedings of the Royal Society of London A* **200**: 331-351.

Pimm, S.L. (1989). "Theories of predicting success and impact of introduced species". In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M.. Chichester. John Wiley and Sons. 351-368.

Pliny, the elder. (AD77). *Naturalis Historia.* From: *Pliny. Natural history (1938), with an English translation in 10 volumes.* Volume IV, Libri XII-XVI. London, William Heinemann Ltd.

Pollard, E. (1979). "A national scheme for monitoring the abundance of butterflies: the first three years." *British Entomological and Natural History Society, Proceedings and Transactions* **12**: 77-90.

Prach, K. (1994). "Seasonal dynamics of *Impatiens glandulifera* in two riparian habitats in central England". In *Ecology and Management of Invasive Riverside Plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 127-133.

Prentice, C.I., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A. & Solomon, A.M. (1992). "A global biome model based on plant physiology and dominance, soil properties and climate." *Journal of Biogeography* **19**: 117-134.

Prince, S.D. & Carter, R.N. (1985) "The geographical-distribution of prickly lettuce (*Lactuca serriola*) .3. Its performance in transplant sites beyond its distribution limit in Britain". *Journal of Ecology* **73** (1): 49-64

Prowse, A. (1998) "Patterns of early growth and mortality in *Impatiens glandulifera*". In *Plant invasions: ecological mechanisms and human responses.* **Eds**: Starfinger, U., Edwards, K., Kowarik, I. & Williamson, M. Leiden. Backhuys Publishers. 245-252.

Pysek, P. (1991). "*Heracleum mantegazzianum* in the Czech Republic: dynamics of spreading from the historical perspective." *Folia Geobotanica et Phytotaxonomica* **26**: 439-454.

Pysek, P. (1994). "Ecological aspects of invasion by *Heracleum mantegazzianum* in the Czech Republic". In: *Ecology and management of invasive riverside plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 45-54.

Pysek, P., Kopecky, M., Jarosik, V. & Kotkova, P. (1998). "The role of human density and climate in the spread of Heracleum mantegazzianum in the Central European landscape." *Diversity and Distributions* **4**: 9-16.

Pysek, P. & Prach, K. (1993). "Plant invasions and the role of riparian habitats: a comparison of four species alien to central Europe." *Journal of Biogeography* **20**: 413-420.

Pysek, P. & Pysek, A. (1995). "Invasion by *Heracleum mantegazzianum* in different habitats in the Czech Republic." *Journal of Vegetation Science* **6**: 711-718.

Ramakrishnan, P.S. (1991). *Ecology of biological invasions in the tropics.* Proceedings of an international workshop. Nainitel, India.

Ratcliffe, D.A. (1968). "An ecological account of Atlantic bryophtes in the British Isles." *New Phytologist* **67**: 365-439.

Ratcliffe, D.A. (1977). *A nature conservation review.* Cambridge. Cambridge University Press.

Ray, J. (1686). *Historia Plantarum.* In Raven, C.E. *"John Ray, naturalist. His life and works."* (1942). Cambridge, Cambridge University Press.

Reeves, S.A. & Usher, M.B. (1989). "Application of a diffusion model to the spread of an invasive species: The coypu in Great Britain." *Ecological Modelling* **47**: 217-232.

Rejmanek, M. (1989). "Invasibility of plant communities". In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M.. Chichester. John Wiley and Sons. 369-388.

Reynolds, S.C. (1996). "Alien plants at ports and in coastal habitats on the east coast of Ireland." *Watsonia* **21**: 53-61.

Richards, C., Haro, R.J., Johnson, L.B. & Host, G.E. (1997). "Catchment and reach-scale properties as indicators of macroinvertebrate species traits." *Freshwater Biology* **37**: 219-230.

Roblin, L. (1994). Alien invasive weeds - an example of National Rivers Authority sponsered research. In: *Ecology and management of invasive riverside plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 189-193.

Rodwell, J.S., Ed. (1991). *British plant communities. Volume 1: Woodlands and scrub.* Cambridge, Cambridge University Press.

Rodwell, J.S., Ed. (1991). *British plant communities. Volume 2: Mires and heaths.* Cambridge, Cambridge University Press.

Rodwell, J.S., Ed. (1992). *British plant communities. Volume 3: Grassland and montane communities.* Cambridge, Cambridge University Press.

Rodwell, J.S., Ed. (1995). *British plant communities. Volume 4: Aquatic communities, swamps and tall-herb fens.* Cambridge, Cambridge University Press.

Rodwell, J.S., Ed. (1999). *British plant communities. Volume 5: Maritime and weed communities and vegetation of open habitats.* Cambridge, Cambridge University Press.

Rorison, I.H. (1981). "Plant growth in response to variation in temperature: field and laboratory studies." In: *Plants and the atmospheric environment.* **Eds**: Grace, J., Ford, E.D. & Jarvis, P.G. British Ecological Society. 313-332.

Rubec, C.D.A. & Lee, G.O. (1997). *Conserving vitality and diversity. Proceedings of the world conservation congress workshop on alien invasive species.* Canada. Canada WS.

Rubow, T. (1979). "Bekaempelse af Kaempe-bjornelko (*Heracleum mantegazzianum).*" *Statens Plantaevlsforsog Meddelelser* **81**: 46-49.

Rushton, S.P., Hill, D. & Carter, S.P. (1994). "The abundance of river corridor birds in relation to their habitats: a modelling approach." *Journal of Applied Ecology* **31**: 313-328.

Salisbury, E. (1961). *Weeds and aliens*. London. Collins.

Scott, R. (1988). *A review of Japanese knotweed control*. Cumbria. NCC/NERC.

Sculthorpe, C.D. (1967). *The biology of aquatic vascular plants*. London. Edward Arnold.

Shumova, E.M. (1972), from Tiley *et al.* (1996). "Dormancy in seeds of giant hogweed." *Doklady Tskha (in Russian)* **180** (2): 235-243.

Simberloff, D. (1986). "Introduced insects: a biogeographic and systematic perspective." In: *Ecology of biological invasions: an Australian perspective*. **Eds**: Groves, R.H. & Burdon, J.J. Canberra. Australian Academy of Science.

Simberloff, D. (1989). "Which insect introductions succeed and which fail?" In: *Biological invasions: a global perspective*. **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. Wiley & Sons. 61-83.

Sindel, B.M. & Michael, P.W. (1992). "Spread and potential distribution of *Senecio madagascariensis* Poir. (fireweed) in Australia." *Australian Journal of Ecology* **17**: 21-26.

Smith, C.W. (1985). "Impact of alien plants on Hawaii's native biota". In: *Hawaii's terrestrial ecosystems: preservation and management*. **Eds**: Stone, C.P. & Scott, J.M. Honalulu. University of Hawaii Press.

Smith, D.B. & Francis, E.A. (1967). *Geology of the country between Durham and West Hartlepool*. London. Mem. Geol. Soc. HMSO.

Smith, K. (1967). "A note on minimum screen temperatures in the Houghall frost hollow." *Met. Mag.* **99**: 45-49.

Smith, K. (1970). Weather and climate. In *Durham county and city with Teeside*. **Ed**.Dewdney, J.C.. Durham.

Smith, P.A. (1994). "Autocorrelation in logistic regression modelling of species' distributions." *Global Ecology and Biogeography Letters* **4**: 47-61.

Sommier, S. & Levier, E. (1895). "Decas umbelliferum novarum caucasi." *Nouvo Giornale Botanico Italiano* **11**(2): 79-81.

Sousa, W.P. (1984). "The role of disturbance in natural communities." *Annual Review of Ecology and Systematics* **15**: 353-391.

Stace, C. (1991). *New Flora of the British Isles*. Cambridge. Cambridge University Press.

Stone, C.P., Smith, C.W. & Tunison, J.T., Eds. (1992). *Alien plant invasions in native ecosystems of Hawaii: management and research*. Honalulu, Hawaii Univ. Press.

Stowe, T.J., Newton, A.V., Green, R.E. & Mayes, E. (1993). "The decline of the corncrake *Crex crex* in Britain and Ireland in relation to habitat." *Journal of Applied Ecology* **30**: 53-62.

Tabachnik, B.G. & Fidell, L.S. (1996). *Using multivariate statistics. 3rd Edition.* New York, USA. Harper Collins.

Tansley A.G. (1939) *The British Isles and their vegetation.* Cambridge, Cambridge University Press.

Thompson K., Hodgson J.G. & Rich T.C.G. (1995) "Native and alien invasive plants: more of the same? " *Ecography* **18:** 390-402.

Tiley, G.E.D., Dodd, F.S. & Wade, P.M. (1996). "Biological flora of the British Isles: *Heracleum mantegazzianum* Sommier and Lever." *Journal of Ecology* **84** (2): 297-320.

Tiley, G.E.D. & Philp, B. (1992). "Strategy for the control of Giant Hogweed (*Heracleum mantegazzianum*) on the River Ayr in Scotland." *Aspects of Applied Biology* **29**: 463-466.

Tiley, G.E.D. & Philp, B. (1994). "*Heracleum mantegazzianum* (Giant Hogweed) and its control in Scotland". In: *Ecology and management of invasive riverside plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 101-109.

Toner, M. & Keddy, P. (1997). "River hydrology and riparian wetlands: a predictive model for ecological assembly." *Ecological Applications* **7**(1): 236-246.

Trewick, S. & Wade, P.M. (1986). *The distribution and dispersal of two alien species of Impatiens, waterway weeds in the British Isles.* Proceedings 7th International European Weed Research Society./Association of Applied Biologists Symposium on Aquatic Weeds. Loughborough. Loughborough University of Techology.

Upton, G.J.G. & Fingleton, B. (1985). *Spatial data analysis by example. Volume I: point pattern and quantitative data.* Chichester. John Wiley & Sons.

Usher, M.B. (1986) "Invasibility and wildlife conservation: invasive species on nature reserves." *Philosophical Transactions of the Royal Society of London* **B314**: 695-710.

van den Bosch, F., Metz, J.A.J. & Diekmann, O. (1990). "The velocity of spatial population expansion." *Journal of Mathematical Biology* **28**: 529-565.

Vitousek, P.M. (1992). "Effects of alien plants on native ecosystems". In: *Alien plant invasions in native ecosystems of Hawaii: management and research.* **Eds**: Stone, C.P., Smith, C.W. & Tunison, J.T. Honalulu. Hawaii Univ. Press.

Vogt Andersen, U. (1994). "Sheep grazing as a method of controlling *Heracleum mantegazzianum.*" In: *Ecology and management of invasive riverside plants.* **Eds**: Brock, J.H., Child, L.E., de Waal, L.C. & Wade, P.M. Chichester. John Wiley and Sons. 77-91.

Vogt Andersen, U. & Calov, B. (1996). "Long term effects of sheep grazing on giant hogweed, (*Heracleum mantegazzianum*)". *Hydrobiologia* **340**: 277-284.

Wadsworth, R.A., Collingham, Y.C., Willis, S.G., Huntley, B. & Hulme, P.E. (in press). "Simulating the spread and management of alien riparian weeds: are they out of control?" *Journal of Applied Ecology.*

Wadsworth, R.A., Swetnam, R.D. & Willis, S.G. (in prep.). "Floating or sinking: modelling the different mechanisms of seed dispersal using *Impatiens glandulifera* Royle and *Heracleum mantegazzianum* Sommier & Levier as examples."

Webb, D.A. (1985). "What are the criteria for presuming native status?" *Watsonia* **15**: 231-236.

Whitton, B.A., Boulton, P.N.G., Clegg, E.M., Gemmell, J.J., Graham, G.G., Gustar, R. & Moorhouse, T.P. (1998). "Long-term changes in macrophytes of British rivers: 1. River Wear." *The Science of the Total Environment* **210**: 411-426.

Whitton, B.A. & Buckmaster, R.C. (1970). "Macrophytes of the River Wear." *The Naturalist* **July-Sept.**: 97-116.

Whitton, B.A. & Dalpra, M. (1968). "Floristic changes in the River Tees." *Hydrobiologia*: **32**: 545-550.

Williamson, J.A. & Forbes, J.C. (1982). *Giant hogweed (Heracleum mantegazzianum): its spread and control with glyphosphate on amenity areas.* Weeds, Proceedings of the 1982 British Crop Protection Conference.

Williamson, M. (1989) "Mathematical models of invasions". In: *Biological invasions: a global perspective.* **Eds**: Drake, J.A., Mooney, H.A., di Castris, F., Groves, R.H., Kruger, F.J., Rejmanek, M. & Williamson, M. Chichester. John Wiley and Sons. 329-350

Willis, S.G. (1995). *"An investigation into the changes in composition and distribution of the vegetation of Widdybank Fell, Upper Teesdale, during the last twenty years".* M.Sc. Thesis. University of Durham.

Willis, S.G., Hulme, P.E. & Huntley, B. (1997). "Habitat suitability and the distribution of alien weeds of riparian ecosystems." In: *Species dispersal and land-use processes: Proceedings of the 6th annual IALE (UK) conference.* University of Ulster, Coleraine. IALE (UK).

Woodward, F.I. (1987). *Climate and plant distribution.* Cambridge. Cambridge University Press.

Woodward, F.I. (1997) "Life at the edge: a 14-year study of a *Verbena officinalis* population's interactions with climate". *Journal of Ecology* **85** (6): 899-906

Wyse Jackson, M. (1989). "Observations on the Irish distribution of a plant with serious public health implications: Giant Hogweed (*Heracleum mantegazzianum*)." *Bulletin of the Irish Biogeographical Society* **12**: 94-112.

Appendix I – RCS Derived Data

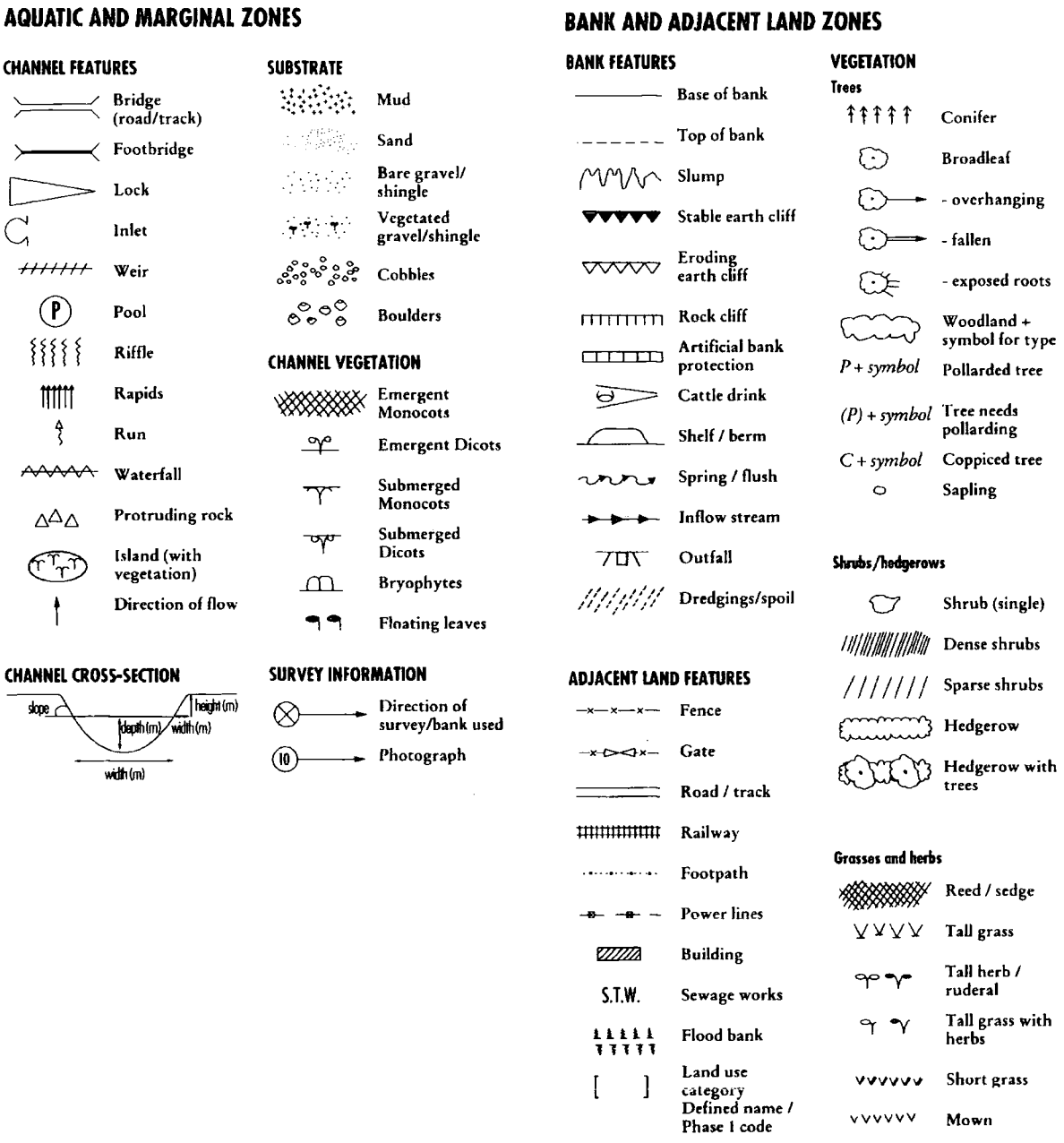

Figure AI.1 Standard habitat and riverine feature recording symbols used on the RCS section maps.

Table AI.1 Abbreviations used to represent RCS variables in tables etc. throughout Chapters 2-4. Composite variables are explained in Appendix Table AI.7.

Variable	Code
Deciduous woodland (bankside)	Decid_ba
Deciduous wooldand (land-use)	Decid_lu
Coniferous woodland (bankside)	Conif_ba
Coniferous woodland (land-use)	Conif_lu
Total Woodland	Tot_wood
Mixed woodland	Mixed
Scrub (bankside)	Scrub
Roads	Roads
Bridges	Bridge
Road or bridge	Ro_or_br
Stream	Stream
All Residential land-use	Tot_res
All Urban Areas	Tot_urb
Derelict areas	Derelict
Hard cliffs	Hard_cls
Soft/slumping riverbank	Soft_slu
Amenity/mown/garden areas	Am/mo/ga
Short/mown grassland	Sh_mo
Arable farmland	Arable
All Farming land-uses	Farming
Grass/herb vegetation	Gr/he
Herb/ruderal vegetation	He/ru
Silt river/bank substrate	Silt
Mud river/bank substrate	Mud
Sand river/bank substrate	Sand
Shingle river/bank substrate	Shingle
Cobble river/bank substrate	Cobbles
Boulder river/bank substrate	Bould
Bedrock river/bank substrate	Bed
Artificial Bank	Artific
Maximum river width	Max_wid
Minimum river width	Min_wid
Maximum river depth	Max_dep
Minimum river depth	Min_dep
Maximum bank slope	Max_slop
Minimum bank slope depth	Min_slop
Grassland (surrounding land-use)	Grass
Grazed bankside	Grazed
Unmanaged bankside	Unman
Heracleum mantegazzianum	Herac
Impatiens glandulifera	Impat
Fallopia japonica	Fallop
Altitude	Altitude

Table AI.2 Riverine morphology features recorded from the River Corridor Surveys. All features recorded as simple presence/absence apart from ripples which was recorded on an abundance scale from 1-5 and pools which were recorded as numbers of pools.

Morphological Features
Cobble Bank
Cobble Shoal
Mudbank
Protruding Rocks
Protruding Stones
Sand
Sandbank
Shelves and Berms
Riffles
Pools

Table AI.3 Bankside features recorded from the River Corridor Surveys. All features were recorded on a scale of 1-10 (see text for details).

Bankside Features		
Coniferous Woods	Tall Grass	Isolated Trees
Hedge	Short Grass	Scrub
Mud	Mown Grass	Reedbeds/Sedges
Sand	Slumping Bank	Herbs/Ruderal
Bare Shingle	Solid Earth Cliff	Deciduous Woodland
Vegetated Shingle	Soft Earth Cliff	Sparse Herb/Ruderal
Cobbles	Artificial Bank	Mixed Grass/Herbs
Boulders	Rock Cliff	Sparse Grass/Herbs
Inlet	Cattle Drink	Inflow Stream
Weir	Spring/Flush	Outfall
Islands	Bridges	Dredgings

Table AI.4 Surrounding land-use categories recorded from the River Corridor Surveys; recorded as simple presence/absence.

Surrounding Land-use	
Deciduous Woodland	Arable Farmland
Coniferous Woodland	Grassland (Improved)
Scrub	Grassland (Semi-improved
Urban Areas	Grassland (Unimproved)
Industrial Areas	Meadow
Roads etc.	Other Ruderal Habitats
Derelict/Building Construction	Nature Reserve/ Natural Habitats
Amenity Grassland	Planted Exotics
Mixed Woodland	Sewage Works
Ungrazed Rough Grass	Non-housing Urban Site
Pond/Lake	Gardens
Improved Meadow	Farm Buildings
Wet Grassland	

Table AI.5 River bed substrate categories recorded from the River Corridor Surveys. Substrates were recored as 0, 1 or 2; 1 representing occurrence but not dominance and 2 representing dominance.

River Bed Substrates	
Silt	Pebbles
Mud	Cobbles
Sand	Boulders
Gravel	Bedrock

Table AI.6 Macrophyte species recorded in the River Corridor Surveys

Algae etc.	Dicotyledonous spp.(cont)	Monocotyledonous spp.
Hildenbrandia spp.	Eupatorium cannibium	Agrostis stolonifera
Lemanea spp.	Filipendula ulmaria	Alisma-plantago aquatica
Vaucheria spp.	*Heracleum* mantegazzianum	Alopecurus geniculatus
Enteromorpha spp.	*Impatiens* glandulifera	Anthoxanthum odoratum
Cladophera glomerata	Lotus uliginosus	Carex acuta
Chlorophyceae spp.	Lysimachia vulgaris	Carex aquatilis
Fucus vesiculosus	Mentha aquatica	Carex binervis
	Mimulus cupreus	Carex flacca
Mosses and Hepatics	Mimulus guttatus x M.luteus	Carex hirta
Conocephalus conicum	Mimulus guttatus	Carex otrubaea
Lunularia cruciata	Myosotis scorpioides	Carex lepidocarpa
Marchiantia polymorpha	Myosoton aquaticum	Carex nigra
Nardia compressa	Nasturcium officinale agg.	Carex pendula
Pellia endiviifolia	Oenanthe crocata	Carex rostrata
Scapania undulata	Petasites hybridus	Catabrosa aquatica
Amblystegium fluviatile	Polygonum amphibium	Tritonia x crocosmiflora
Amblystegium riparium	Potentilla palustris	Deschampsia cespitosa
Brachythecium plumosum	Pulicaria dysenterica	Eleocharis palustris
Brachythecium rivulare	Ranunculus aquatilis	Elodea canadensis
Brachythecium rutabulum	Ranunculus circinatus	Elodea nuttallii
Calliergon cuspidatum	Ranunculus flammula	Festuca aruninacea
Cinclidotus fontinaloides	Ranunculus fluitans	Glyceria declinata
Fontinalis antipyretica	Ranunculus penicillatus	Glyceria fluitans
Rhynchostegium riparioides	Ranunculus repens	Glyceria plicata
Schistidium alpicola	Ranunculus trichophyllus	Iris pseudacorus
	Ranunculus sceleratus	Juncus acutiflorus
Dicotyledonous spp.	Ranunculus vert.	Juncus articulatus
Equisetum arvensis	Reynoutria japonica	Juncus bufonius
Equisetum fluviatile	Reynoutria sachelinensis	Juncus tenuis
Equisetum hyemale	Rorippa amphibia	Juncus effusus
Equisetum palustre	Rorippa sylvestris	Juncus inflexus
Filicopsida spp.	Sagina procumbens	Juncus squarrosus
Angelica sylvestris	Scrophularia auriculata	Lemna minor
Apium inundatum	Senecio aquaticus	Phalaris arundinacea
Apium nodiflorum	Solanum dulcamara	Phragmites australis
Callitriche hamulata	Stachys palustris	Poa stolonifera ?
Callitriche obtusangula	Symphytum officinale	Poa trivialis
Callitriche platycarpa	Thalictrum flavum	Potamageton crispus
Callitriche stagnalis	Tussilago farfara	Potamageton friesii ?
Callitriche	Veronica beccabunga	Potamageton natans
Cardamine amara	Veronica scutellata	Potamageton pectinatus
Epilobium hirsutum		Scirpus maritimus
		Sparganium erecta
		Typha latifolia

Table AI.7 The amalgamated variables created

Composite variable	Code	Original constituent variables (lu and ba representing land-use and bankside variables respectively)
Amenity/Mown Grass/Garderns	Am_mo_ga	Amenity Grassland, Mown Grassland, Gardens (lu)
Derelict/Ruderal	Derelict	Derelict and Other Ruderal (lu)
Disturbance	Disturb	Aritificial banks, Cattle drinks, Dredgings/spoil etc., Weirs (ba)
Farming	Farming	Arable land, Improved Grassland, Semi-improved Grassland, Unimproved Grassland, Meadows and Improved Meadows (lu)
Hard Cliffs	Hard_cls	Rock Cliffs and Solid Earth Cliffs (ba)
Mixed/Conifer Wood	Mixed	Mixed Woodland and Coniferous Woodland (lu)
Road/Bridge	Road/br	Roads beside or Bridges over river (lu)
Total Herb/Ruderal	He/ru	Herb/Ruderal and Sparse Herb/Ruderal Vegetation (ba)
Total Mixed Grass/Herbs	Gr/he	Mixed Grass/Herbs and Sparse Mixed Grass/Herbs (ba)
Total Residential	Tot_res	Residential Areas, Non-residential Urban Areas (lu)
Wilderness areas	Wild	Ponds etc, Natural Areas/Nature Reserves, Scrub, Rough Grassland and Wet Grassland (lu)
Total Woodland	Tot_wood	Deciduous Woodland, Coniferous Woodland (ba)

Appendix II – Exploratory Statistics

i) Exploratory Statistics for Quantitative RCS Variables

Almost all of the RCS variables examined showed a marked degree of skew, suggesting non-normal distributions. Additionally many of the variables had values of kurtosis which run contrary to the expectations for a normal distribution. These findings apply equally to both the Tees and the Wear variables. Kolmogorov-Smirnoff tests of all the variables indicated that no variables were normally distributed. Taking logarithms of the data was found not to aid in normalising the data. Again Kolmogorov-Smirnoff tests on the transformed variables confirmed a lack of normality. Mann-Whitney U-tests were therefore used to examine the relationship between the study species and these variables (Tables AII2-AII4).

Table AII.1: Summary statistics for the quantitative amalgamated variables of the Wear and the Tees catchment databases. N=365 for the Wear data and N=871 for the Tees data. Range column differentiates between continuous and limited ranges of variables.

		Wear Catchment				Tees Catchment			
Variable	**Range**	**Min.**	**Max.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Mean**	**Std. Dev.**
Gr/He	cont.	0	32	8.123	7.272	0	40	5.619	6.891
Altitude	cont.	5	340	82.781	69.432	0	530	78.094	88.159
Artific.	0-20	0	20	1.570	3.280	0	20	0.953	3.097
Bridge	cont.	0	4	0.564	0.770	0	10	0.597	0.899
Conif_ba	0-20	0	16	3.342	4.032	0	20	0.603	2.080
Decid_ba	0-20	0	20	3.685	4.296	0	20	4.266	5.649
Max_dep	cont.	0	8	2.378	1.348	0	6	2.192	1.363
Min_dep	cont.	0	8	1.611	1.191	0	4	1.550	1.013
Sh_mo	0-20	0	4	1.329	0.987	0	5	1.437	1.050
Islands	cont.	0	9	0.622	1.240	0	8	0.238	0.681
Max_ht	cont.	0	59	5.414	7.110	0	50	3.984	4.793
Max_slop	0-90.	0	90	71.863	20.820	0	90	67.784	24.614
Min_ht	cont.	0	8	1.408	1.220	0	40	1.400	1.753
Min_slop	0-90.	0	90	36.699	16.219	0	90	38.598	17.418
Reeds	0-20	0	13	1.885	2.328	0	20	4.481	5.317
Scrub_ba	0-20	0	17	3.334	2.668	0	20	3.341	3.817
Sh_mo	0-20	0	19	2.496	3.859	0	20	2.998	5.159
Soft_slu	cont.	0	25	3.134	4.430	0	20	1.759	2.924
Stream	cont.	0	5	0.488	0.794	0	5	0.439	0.798
He_ru	cont.	0	19	3.296	3.935	0	37	5.778	7.103
Weirs	cont.	0	6	0.279	0.670	0	5	0.126	0.447
Max_wid	cont.	0	9	2.718	1.869	0	9	2.629	2.099
Min_wid	cont.	0	9	2.553	1.819	0	9	2.389	2.006

ii) Exploratory Statistics for Binary and Categorical RCS Variables

For the remaining binary variables, Chi^2 contingency tests were used to determine effects of these variables on the species' distributions. Categorical variables were examined using Chi^2 tests of heterogeneity, with categories being combined, where appropriate, when expected values were less than five (Tables AII2-AII4).

Table AII.2 Exploratory summary statistics examining the relationships between the occurrence of *Impatiens* and the RCS variables along the two study catchments. The symbols M and C in the test column indicate the use of Mann-Whitney U-tests or χ^2 tests respectively. The columns "mean abs" and "mean pre" show the mean Mann-Whitney rank for cases where a variable was respectively absent or present

		Tees					Wear				
Impatiens **Variable**	**Test**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**
Grass/Herb	M	64980		0.001	475	365	8279		0.001	139	197
Altitude	M	52364		0	497	325	4525.5		0	270	156
Am/Mo/Gdn	C	6.773	1	0.01			0.001	1	ns		
Arable	C	20.942	1	0.001			6.076	1	0.05		
Artificial	M	80234.5		0.01	425	457	9422.5		0.001	214	173
Bedrock	C	0.024	1	ns			4.670	1	0.05		
Boulders	C	46.089	1	0.001			0.088	1	ns		
Bridges	C	21.952	3	0.001			4.914	2	ns		
Cobbles	C	144.311	1	0.001			0.957	1	ns		
Conifer (ba)	M	83225.5		ns	430	447	10784		ns	168	188
Conifer (lu)	C	0.024	1	ns			0.027	1	ns		
Decid (ba)	M	58643		0	386	527	10054		0.01	160	190
Decid (lu)	C	46.043	1	0.001			7.939	1	0.01		
Max Depth	M	64916.5		0.001	397	507	8108.5		0.001	137	197
Min Depth	M	57580		0	384	531	9039		0.001	148	194
Derelict	C	0.852	1	ns			1.844	1	ns		
Disturb	C	2.052	1	ns			11.629	1	0.001		
Farming	C	45.183	1	0.001			0.018	1	ns		
Grazed	C	1.784	1	ns			3.316	1	ns		
Unman	C	85.673	1	0.001			48.163	1	0.001		
Grassland	M	82274		ns	444	422	8762		0.001	221	171
Hard Cliffs	M	84548		ns	432	443	11464		ns	176	185
Heracleum	C	84.857	1	0.001			6.265	1	ns		
Human	C	5.020	1	0.05	446	417	0.677	1	ns	192	180
Impatiens	C	-	-	-			-	-	-		
Islands	C	13.096	3	0.01	422	461	8.433	3	0.05	164	189
Land-use	C	74.932	14	0.001			56.811	14	0.001		
Max Ht	M	59366		0	387	525	9986.5		0.01	159	191
Max Slope	M	80750.5		ns	425	455	11090		ns	195	179
Min Ht	M	71819.5		0.001	410	484	10186.5		0.05	161	190
Min Slope	M	68571		0.001	468	377	10582		ns	166	188
Mixed	C	0.000	1	ns			0.878	1	ns		
Mud	C	7.337	1	0.01			2.215	1	ns		
Reed	M	86346		ns	435	437	11142		ns	172	186
Fallopia	C	3.415	1	ns			8.663	1	0.01		
Road/Br	C	18.190	1	0.001			1.276	1	ns		
Roads	C	0.676	1	ns			0.193	1	ns		
Sand	C	54.385	1	0.001			2.557	1	ns		
Scrub (ba)	M	86143.5		ns	435	438	11840.5		ns	180	184
Scrub (lu)	C	2.597	1	ns			0.763	1	ns		
Shingle	C	77.047	1	0.001			0.979	1	ns		

Table AII.2 (cont.) Exploratory summary statistics examining the relationships between the occurrence of *Impatiens* and the RCS variables along the two study catchments. The symbols M and C in the test column indicate the use of Mann-Whitney U-tests or χ^2 tests respectively. The columns "mean abs" and "mean pre" show the mean Mann-Whitney rank for cases where a variable was respectively absent or present

		Tees					Wear				
Impatiens **Variable**	**Test**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**
Short/Mown	M	77079		0.01	419	467	11082		ns	195	179
Silt	C	95.26	2	0.001			22.902	2	0.001		
Soft/slump	M	69879.5		0.001	406	491	8842.5		0.001	220	171
Stream	C	8.564	3	0.05			0.68	2	ns		
Substrate	C	49.764	5	0.001			16.277	5	0.01		
Herb/Rud	M	56601.5		0	383	534	6386.5		0	117	204
Tot_Res	C	6.437	1	0.05			1.020	1	ns		
Tot_Wood	C	66.908	1	0.001			18.158	1	0.001		
Urban	C	12.428	1	0.001			1.874	1	ns		
Weir	C	4.701	2	ns			2.568	1	ns		
Wid_max	M	42058		0	357	581	11171.5		ns	194	180
Wid_min	M	46709.5		0	365	566	11445		ns	190	181
Wild	C	2.386	1	ns			0.805	1	ns		

Table AII.3 Exploratory summary statistics examining the relationships between the occurrence of *Heracleum* and the RCS variables along the two study catchments. The symbols M and C in the test column indicate the use of Mann-Whitney U-tests or χ^2 tests respectively. The columns "mean abs" and "mean pre" show the mean Mann-Whitney rank for cases where a variable was respectively absent or present

		Tees					Wear				
Heracleum **Variable**	**Test**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**
Grass/Herb	M	27837.5		0.05	442	377	8070.5		ns	179	204
Altitude	M	8360.5		0	467	143	1100.5		0	210	49
Am/Mo/Gdn	C	3.666	1	ns			0.046	1	ns		
Arable	C	42.358	1	0.001			0.017	1	ns		
Artificial	M	32047.5		ns	437	428	8510		ns	180	197
Bedrock	C	0.360	1	ns			3.807	1	ns		
Boulders	C	18.679	1	0.001			40.156	1	0.001		
Bridges	C	32.736	2	0.001			2.534	2	ns		
Cobbles	C	40.933	1	0.001			0.028	1	ns		
Conifer (ba)	M	31852.5		ns	435	446	7463.5		0.01	177	214
Conifer (lu)	C	0.012	1	ns			0.145	1	ns		
Decid (ba)	M	18907.5		0	418	602	7581.5		0.01	177	212
Decid (lu)	C	30.281	1	0.001			24.023	1	0.001		
Max Depth	M	18498.5		0	418	607	8231.5		ns	179	202
Min Depth	M	14989		0	414	649	7055		0.001	175	221
Derelict	C	2.607	1	ns			14.971	1	0.001		
Disturb	C	6.919	1	0.01			8.939	1	0.01		
Farming	C	20.539	1	0.001			5.325	1	0.05		
Grazed	C	0.031	1	ns			3.015	1	ns		
Unmanage	C	74.801	1	0.001			0.131	1	ns		
Grassland	M	23650.5		0.001	447	327	8792.5		ns	185	173
Hard Cliffs	M	28820		0.01	441	389	7882		0.01	178	207
Heracleum	C	-	-	-			-	-	-		
Human	C	3.601	1	ns	439	403	0.001	1	ns	183	184
Impatiens	C	166.234		0.001			75.988		0.001		
Islands	C	0.114	1	ns	437	425	3.513	2	ns	179	203
Land-use	C	86.366	14	0.001			50.458	14	0.001		
Max Ht	M	31989.5		ns	437	427	6151.5		0.001	172	235
Max Slope	M	26374		0.01	444	360	9253.5		ns	183	181
Min Ht	M	31387.5		ns	434	452	6576		0.001	174	228
Min Slope	M	14841.5		0	459	221	7799.5		0.05	178	209
Mixed	C	7.801	1	0.01			6.139	1	0.05		
Mud	C	0.600	1	ns			33.930	1	0.001		
Reed	M	28100.5		0.05	430	491	7814		0.05	178	208
Fallopia	C	8.847	1	0.01			20.086	1	0.001		
Road/Br	C	44.422	1	0.001	449	315	0.955	1	ns	181	193
Roads	C	0.033	1	ns			22.359	1	0.001		
Sand	C	19.129	1	0.001			1.177	1	ns		
Scrub (ba)	M	28111.5		0.05	430	491	8770.5		ns	181	193
Scrub (lu)	C	1.759	1	ns			0.103	1	ns		
Shingle	C	50.500	1	0.001			1.854	1	ns		

Table AII.3 (cont.) Exploratory summary statistics examining the relationships between the occurrence of *Heracleum* and the RCS variables along the two study catchments. The symbols M and C in the test column indicate the use of Mann-Whitney U-tests or χ^2 tests respectively. The columns "mean abs" and "mean pre" show the mean Mann-Whitney rank for cases where a variable was respectively absent or present

		Tees					Wear				
Heracleum **Variable**	**Test**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**
Short/Mown	M	29988		ns	433	469	6501.5		0.001	173	230
Silt	C	68.264	2	0.001			4.423	1	0.05		
Soft/slump	M	21923		0.001	422	566	7242		0.01	190	148
Stream	C	1.631	2	ns			0.66	2	ns		
Substrate	C	11.347	5	0.05			21.999	5	0.001		
He/Ru	M	23882		0.001	425	542	8709		ns	181	194
Tot_res	C	3.427	1	ns			0.221	1	ns		
Tot_wood	C	66.628	1	0.001			24.601	1	0.001		
Urban	C	16.567	1	0.001			0.145	1	ns		
Weir	C	0.777	1	ns			4.841	1	0.05		
Max Width	M	7516		0	404	739	2575		0	160	293
Min Width	M	7753.5		0	404	737	2312.5		0	160	297
Wild	C	5.568	1	0.05			0.003	1	ns		

Table AII.4 Exploratory summary statistics examining the relationships between the occurrence of *Fallopia* and the RCS variables along the two study catchments. The symbols M and C in the test column indicate the use of Mann-Whitney U-tests or χ^2 tests respectively. The columns "mean abs" and "mean pre" show the mean Mann-Whitney rank for cases where a variable was respectively absent or present

		Tees					Wear				
Fallopia **Variable**	**Test**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**
Grass/Herb	M	4029		ns	438	317	12783.5		ns	179	192
Altitude	M	2477.5		0.001	440	198	13659.5		ns	183	184
Am/Mo/Gdn	C	3.221	1	ns			0.001	1	ns		
Arable	C	0.078	1	ns			1.455	1	ns		
Artificial	M	3919		0.05	434	564	12805.5		ns	187	174
Bedrock	C	1.454	1	ns			5.810	1	0.05		
Boulders	C	3.923	1	0.05			1.413	1	ns		
Bridges	C	0.0165	1	ns			0.335	2	ns		
Cobbles	C	4.153	1	ns			1.008	1	ns		
Conifer (ba)	M	5190.5		ns	436	406	12287		ns	177	197
Conifer (lu)	C	0.049	1	ns			0.003	1	ns		
Decid (ba)	M	3601.5		0.05	434	588	12472.5		ns	178	195
Decid (lu)	C	4.934	1	0.05			3.509	1	ns		
Max Depth	M	5387		ns	436	421	13711		ns	183	183
Min Depth	M	4269		ns	434	537	11684.5		0.05	175	202
Derelict	C	0.106	1	ns			0.007	1	ns		
Disturb	C	1.503	1	ns			1.195	1	ns		
Farming	C	20.360	1	0.001			5.938	1	0.05		
Grazed	C	5.646	1	0.05			1.851	1	ns		
Unmanage	C	2.899	1	ns			8.568	1	0.01		
Grassland	M	3128.5		0.01	439	248	12109.5		ns	189	168
Hard Cliffs	M	4927		ns	437	386	13172.5		ns	185	178
Heracleum	C	10.489	1	0.01			14.676	1	0.001		
Human	C	0.048	1	ns			1.132	1	ns		
Impatiens	C	7.931		0.01			76.775		0.001		
Islands	C	0.184	1	ns			1.523	3	ns		
Land-use	C	22.655	14	ns			7.933	14	ns		
Max Ht	M	3801		0.05	434	573	13613		ns	183	182
Max Slope	M	3632		0.05	434	586	11186		0.01	193	159
Min Ht	M	4802.5		ns	435	496	13516.5		ns	182	185
Min Slope	M	4282.5		ns	438	336	13006		ns	186	176
Mixed	C	0.457	1	ns			0.267	1	ns		
Mud	C	0.189	1	ns			2.387	1	ns		
Reed	M	5241		ns	436	410	12498		ns	188	171
Fallopia	C		1				-	-	-		
Road/Br	C	0.282	1	ns			0.391	1	ns		
Roads	C	8.303	1	0.01			0.388	1	ns		
Sand	C	0.192	1	ns			1.232	1	ns		
Scrub (ba)	M	4943		ns	435	485	12859		ns	186	175
Scrub (lu)	C	0.141	1	ns			0.004	1	ns		
Shingle	C	1.962	1	ns			6.104	1	0.05		

Table AII.4 (cont.) Exploratory summary statistics examining the relationships between the occurrence of *Fallopia* and the RCS variables along the two study catchments. The symbols M and C in the test column indicate the use of Mann-Whitney U-tests or χ^2 tests respectively. The columns "mean abs" and "mean pre" show the mean Mann-Whitney rank for cases where a variable was respectively absent or present

		Tees					Wear				
Fallopia **Variable**	**Test**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**	**Statistic**	**d.f.**	**p**	**Mean abs**	**Mean pre**
Short/Mo	M	4285.5		ns	438	337	13386		ns	182	186
Silt	C	2.867	1	ns			3.672	2	ns		
Soft/slump	M	4286		ns	434	535	9661.5		0.001	199	145
Stream	C	0.147	1	ns			3.114	2	ns		
Substrate	C	*1.048*	4	ns			8.109	5	ns		
He/Ru	M	3104.5		0.01	433	626	13634.5		ns	183	184
Tot_res	C	10.767	1	0.01			0.903	1	ns		
Tot_wood	C	0.498	1	ns			4.584	1	0.05		
Urban	C	5.271	1	0.05			0.957	1	ns		
Weir	C	0.0001	1	ns			2.028	1	ns		
Max Wid	M	3056.5		0.01	433	630	12804.5		ns	179	192
Min Wid	M	2833		0.001	433	647	12910.5		ns	180	191
Wild	C	0.024	1	ns			2.568	1	ns		

Variable	gr/he	altitude	am_mo	arable	artific	bed	bould	bridge	cobbles	conif_ba	conif_le	decid_ba	decid_lu	max_dep	min_dep	derelict	disturb	farming	grazed	unman	grass	hard_cls	Herac	humen	Impat	island	max_ht	max_slop	min_ht
gr/he																													
altitude	-0.184																												
am_mo		-0.135																											
arable	0.236	-0.231	-0.119																										
artific			0.226	-0.275																									
bed		0.223																											
bould		0.296	-0.149	-0.135		0.338																							
bridge	-0.177		0.199	-0.134	0.261		-0.109																						
cobbles		0.110	-0.104			0.154	0.529																						
conif_ba				-0.219		0.336	0.263		0.197																				
conif_le						0.176	0.103			0.166																			
decid_ba			-0.158	-0.132		0.331	0.144		0.113	0.442																			
decid_lu		-0.180								0.456		0.323																	
max_dep	0.127			0.111	-0.146	0.117	0.233	-0.141	0.230	0.146	0.115	0.106																	
min_dep	0.171				-0.111		0.234	-0.150	0.359	0.207		0.145		0.718															
derelict		0.229		-0.197						0.119		0.117			0.116														
disturb		0.155	0.134	-0.344	0.657			0.267		0.104				-0.108		0.413													
farming		0.113	-0.135	0.310	-0.146					-0.140				0.138			-0.130												
grazed						-0.167				-0.190		-0.133	-0.162					0.222											
unman		-0.138		0.139						0.181	0.121	0.131	0.212	0.159	0.159				-0.408										
grass		0.258	-0.113							-0.172			-0.205					0.601	0.429	-0.287									
hard_cls		-0.175				0.188			0.140	0.230		0.192	0.226					-0.135			-0.115								
Herac		-0.575				-0.105	-0.238			0.142		0.129	0.211		0.172	-0.208	-0.117					0.154							
human			0.446	-0.163	0.239		-0.118	0.632							-0.140		0.215												
Impat	0.234	-0.462		0.222	-0.183	-0.206		-0.115				0.128	0.225	0.256	0.198	-0.138	-0.239		-0.179	0.257	-0.213		0.253						
island							0.122			0.139		0.137		0.124									0.103		0.122				
max_ht	0.186	-0.292				0.146	0.120		0.147	0.115		0.124		0.177	0.176							0.269	0.228		0.130				
max_slop	-0.123	-0.145			0.115				-0.119						-0.209							0.187					0.141		
min_ht	0.209	-0.311		0.192				-0.188						0.215	0.258				-0.153	0.182			0.203		0.121	-0.122	0.426		
min_slop	-0.203	-0.264	0.150	0.113			-0.211	0.176	-0.138						-0.229	-0.125							0.112	0.147				0.199	0.126
mixed										0.201				0.155									0.151		0.109				
mud		-0.421	0.160	0.139		-0.321	-0.395		-0.372	-0.132				-0.127	-0.233	-0.214							0.237	0.135	0.116			0.197	
reeds	0.146	-0.193		0.136		-0.165	-0.280	0.123	-0.234	-0.279	-0.157	-0.126			-0.179	-0.157		0.117		-0.140	0.127		0.113	0.151		-0.117			
Fall						-0.135									0.124								0.225		0.273			-0.154	
road_br	-0.154		0.193	-0.139	0.246			0.874						-0.104	-0.128	-0.110	0.234							0.722					-0.103
road		-0.189	0.126		0.261			0.232									0.156	-0.139			-0.162		0.257	0.344			0.149		0.170
sand	0.106	-0.124		0.150		-0.169			-0.160									0.129							0.105				
scrub_ba	0.188	-0.213		0.246				0.158		-0.311			-0.107			-0.160				0.104				0.163			0.157	0.103	0.172
scrub_lu	-0.154			-0.133				0.108	0.116	-0.159		-0.129									0.123						-0.108	0.103	-0.144
shingle								0.157	-0.296			0.116		-0.135	-0.236	-0.127					0.119			0.144		0.111			-0.104
sh_mo	0.137		0.198	-0.120	0.230	-0.107	-0.117					-0.117							0.435	-0.326	0.216		0.218	0.107			0.135		
silt	0.176								0.178						0.203	0.171		-0.140	-0.118	0.153	-0.147				0.250				
soft_cls			-0.162						-0.227	-0.112				-0.200	-0.316	-0.119		0.174	0.239	-0.215	0.250		-0.153		-0.204		-0.130	0.307	-0.264
stream						0.104						0.172	0.166											-0.115					
he/ru		-0.211		0.208	-0.125	-0.250	-0.163										-0.165	0.123							0.353		-0.128		
tot_res			0.187	-0.162	0.270			0.289					-0.120	-0.109			0.196	-0.137						0.427					
tot_wood						0.239	0.142		0.141	0.487	0.174	0.555	0.481	0.136	0.179		0.111		-0.157	0.216		0.104	0.137		0.214		0.103		
urban	-0.148		0.332	-0.214	0.388		-0.115	0.374					-0.112	-0.149	-0.139		0.325	-0.179						0.384	-0.134				
weir		0.238		-0.222	0.173	0.159		0.244									0.418	-0.116					-0.118	0.214	-0.164		-0.119		
max_wid	0.131	-0.386			0.271	0.212			0.225	0.289	0.111	0.139	0.178	0.172	0.294							0.145	0.487				0.361		0.442
min_wid	0.134	-0.410			0.234	0.155		-0.130	0.236	0.270			0.159	0.144	0.299							0.110	0.507				0.341		0.449
wild						-0.129				-0.106				-0.114							0.174	-0.103		0.109			-0.118		-0.129

Table AII.5 Spearman rank correlation analyses between all RCS variables extracted from the Tees catchment RCS.

Variable	min_slop	mixed	mud	reeds	Fall	road_br	road	sand	scrub_ba	scrub_lu	shingle	sh_mo	silt	soft_cls	stream	he/ru	tot_res	tot_wood	urban	weir	max_wid	min_wid	wild
gr/he																							
altitude																							
am_mo																							
arable																							
artific																							
bed																							
bould																							
bridge																							
cobbles																							
conif_ba																							
conif_la																							
decid_ba																							
decid_lu																							
max_dep																							
min_dep																							
derelict																							
disturb																							
farming																							
grazed																							
unman																							
grass																							
hard_cls																							
Herac																							
human																							
Impat																							
island																							
max_ht																							
max_slop																							
min_ht																							
min_slop																							
mixed																							
mud	0.202																						
reeds	0.150		0.256																				
Fall																							
road_br	0.170			0.139																			
road			0.112			0.476																	
sand			0.128																				
scrub_ba	0.134		0.255	0.270		0.202																	
scrub_lu								0.139															
shingle			0.112	0.262	-0.107	0.191			0.123														
sh_mo	-0.138		0.121	0.104			0.119																
silt		-0.122	-0.151					-0.203		-0.160	-0.232												
soft_cls			0.156		-0.240		-0.139	0.198		0.140	0.330		-0.205										
stream																							
he/ru	0.212	0.137		0.258				0.160															
tot_res		-0.112				0.298	0.332	-0.122	0.124			0.123		-0.117	-0.127	-0.123							
tot_wood		0.143						-0.104					0.148		0.181								
urban			0.169			0.340	0.319	-0.135				0.108		-0.111	-0.109		0.684						
weir						0.237	0.103				0.143					-0.133	0.227	0.107	0.191				
max_wid							0.191	-0.110	0.131			0.190		-0.215		-0.243	0.117	0.199	0.112				
min_wid							0.180		0.127			0.193		-0.225		-0.212	0.109	0.172			0.944		
wild								0.193		0.710	0.114		-0.190	0.138									

Table AII.5 (cont.) Spearman rank correlation analyses between all RCS variables extracted from the Tees catchment RCS.

Variable	gr/he	altitude	am_mo	arable	artific	bed	bould	bridge	cobbles	conif_ba	conif_la	decid_ba	decid_lu	max_dep	min_dep	derelict	disturb	farming	grazed	unman	grass	hard_cls	Herac	human	Impat	island	max_ht	max_slop	min_ht
gr/he																													
altitude																													
am_mo		-0.126																											
arable	0.130	-0.192	-0.102																										
artific		-0.188	0.256	-0.136																									
bed		0.364		-0.241																									
bould	-0.187	0.325		-0.210		0.488																							
bridge			0.203		0.234	-0.158	-0.278																						
cobbles	-0.201	0.233	0.085	-0.074		0.340	0.630	-0.265																					
conif_ba		0.153	0.067			0.216	0.175		0.152																				
conif_la		0.095								0.370																			
decid_ba	-0.163					0.331	0.423	-0.237	0.349	0.345	0.111																		
decid_lu	-0.101					0.165	0.260	-0.178	0.206	0.160		0.555																	
max_dep	-0.083		0.096	-0.176	0.109	0.164	0.344	-0.211	0.381	0.109		0.168	0.133																
min_dep	-0.081	-0.167	0.075	-0.072	0.163		0.256	-0.141	0.243		-0.078	0.124		0.752															
derelict		-0.130		-0.125	0.207	-0.075		0.120	-0.069	-0.083		-0.124		0.088	0.122														
disturb		-0.108	0.174	-0.136	0.714			0.210						0.105	0.143	0.412													
farming		0.106	-0.148	0.417	-0.171					0.079		0.182	0.145			-0.184	-0.102												
grazed		0.172		-0.256		0.082	0.243	-0.107	0.175					0.244	0.208														
unman	0.100	-0.119	-0.080	0.259				-0.071		0.083	0.078	0.252	0.201			0.078		0.239	-0.395										
grass		0.205	-0.111	-0.140			0.100	0.073		0.075		0.106	0.078					0.468	0.204										
hard_cls	-0.108	0.247		-0.170		0.487	0.368	-0.129	0.228	0.124		0.194	0.088	0.144					0.106	0.068	0.090								
Herac	-0.077	-0.379	0.067	0.159			0.127	-0.193	0.159			0.226	0.163	0.226	0.296		-0.078			0.143	-0.147	-0.089							
human	0.082	-0.109	0.410		0.253	-0.103	-0.181	0.743	-0.170			-0.133	-0.082	-0.107		0.129	0.228		-0.071			-0.106							
Impat	-0.212	-0.328	0.103	0.132	0.083		0.230	-0.138	0.346			0.282	0.235	0.213	0.299			0.103		0.177			0.373						
island	-0.088	0.115		-0.116		0.247	0.300	-0.157	0.285	0.103		0.246	0.154	0.133	0.088	-0.090			0.093		0.082	0.176		-0.130	0.116				
max_ht	-0.104	0.245				0.340	0.537	-0.106	0.370	0.181	0.085	0.371	0.246	0.135	0.098			0.213	0.077	0.102	0.211	0.290		-0.079	0.266	0.217			
max_slop		0.118				0.193	0.203		0.151	0.104		0.173	0.103		-0.108			0.134		0.067	0.116	0.279	-0.101			0.149	0.363		
min_ht	0.108		-0.072	0.153			0.128		0.087						0.105			0.178		0.155	0.095				0.148		0.469	0.159	
min_slop	0.105	0.222	-0.087		-0.095		-0.134	0.191	-0.119			-0.119		-0.217	-0.226		-0.094	0.078	-0.104	0.100			-0.281	0.117	-0.175		0.085	0.231	0.274
mixed		0.182		-0.137		0.186	0.137	-0.071	0.112	0.578		0.321	0.106	0.133							0.080	0.154	-0.102			0.107	0.143	0.076	
mud	0.225	-0.332		0.210		-0.349	-0.505	0.238	-0.389	-0.237		-0.409	-0.200	-0.245	-0.150	0.090			-0.177	0.104	-0.110	-0.218		0.221	-0.080	-0.229	-0.325	-0.167	
reeds	0.231	-0.139		0.274		-0.295	-0.242	0.187	-0.164	-0.196		-0.363	-0.180	-0.099				0.092				-0.236	0.072	0.114		-0.146	-0.167	-0.108	0.208
Fall		-0.117	0.068		0.084							0.078	0.072					-0.069	-0.079		-0.096		0.121		0.087		0.068	0.075	
road_br	0.088		0.233		0.230	-0.133	-0.245	0.929	-0.252			-0.211	-0.150	-0.192	-0.137	0.103	0.208		-0.095		0.067	-0.120	-0.181	0.800	-0.136	-0.145	-0.082		
road			0.226		0.187			0.163						0.084		0.068	0.158	-0.072						0.247				0.078	
sand	-0.086	-0.091		0.120			0.083	-0.073	0.176	0.077		0.247	0.143			-0.105		0.163					0.127		0.248	0.163	0.167	0.092	
scrub_ba		-0.211		0.245		-0.219	-0.146	0.083	-0.214	-0.130		-0.121		-0.168				0.171	-0.123	0.158	0.099	-0.110	0.072				-0.081		0.105
scrub_lu				-0.073			0.077	-0.090	0.073			0.113	0.182	0.087							0.095			-0.093	0.069	0.099	0.081		
shingle	-0.075		0.098	0.070			0.227	-0.104	0.392			0.229	0.157	0.241	0.163	-0.086		0.123	0.134		0.088		0.161	-0.068	0.233	0.134	0.243	0.202	0.156
sh_mo	-0.172		0.189	-0.166			0.110		0.180	-0.076		-0.069	-0.117	0.211	0.252				0.543	-0.397	0.115				0.105			-0.086	
silt		-0.248			0.102	-0.248	-0.222	0.120	-0.186	-0.196	-0.081	-0.293	-0.225		0.110	0.140	0.118	-0.078	0.079		-0.097	-0.210		0.130	-0.163	-0.132	-0.238	-0.123	
soft_cls		-0.224					0.139	-0.174	0.113			0.186	0.161	0.109	0.164	-0.095			0.249		0.092		0.183	-0.127	0.175	0.160	0.113	0.208	0.080
stream		0.298	-0.075	-0.151	-0.118	0.181	0.183	-0.085	0.143					0.093		-0.079	-0.130		0.128	-0.081		0.138		-0.087	-0.095		0.094		
he/ru	-0.252	-0.356		0.214	0.117	-0.269	-0.171	0.092		-0.169		-0.227	-0.130			0.113		0.068	-0.166	0.167	-0.133	-0.118	0.141	0.097	0.295	-0.105		-0.069	0.157
tot_res		-0.121	0.466	-0.118	0.293			0.142	0.080			0.089			0.074		0.173	-0.117			-0.081		0.067	0.351	0.103				-0.113
tot_wood	-0.113		0.086			0.208	0.296	-0.178	0.251	0.338	0.194	0.815	0.595	0.156	0.116	-0.072		0.198		0.223	0.126	0.119	0.190		0.223	0.186	0.284	0.101	
urban		-0.167	0.626	-0.100	0.291			0.162	0.068	0.089		0.087		0.097	0.102	0.085	0.214	-0.139			-0.091		0.135	0.441	0.136	0.069			-0.081
weir			0.139	-0.098	0.272	0.122		0.135	0.094	0.122							0.415	-0.075						0.125		0.070			
max_wid	-0.086		0.113	-0.159	0.132	0.369	0.641	-0.343	0.577	0.129		0.498	0.272	0.526	0.467		0.069		0.188	0.166		0.254	0.412	-0.167	0.449	0.286	0.400	0.070	
min_wid			0.103	-0.172	0.143	0.372	0.621	-0.355	0.552	0.134	0.074	0.516	0.261	0.482	0.453				0.188	0.172		0.231	0.418	-0.165	0.412	0.274	0.393		0.075
wild		0.257		-0.229		0.197	0.221	-0.116	0.204	0.080		0.118	0.100	0.224	0.089				0.215	-0.113	0.424	0.129		-0.139		0.192	0.150	0.097	

Table AII.6 Spearman rank correlation analyses between all RCS variables extracted from the Wear catchment RCS.

Variable	min_slop	mixed	mud	reeds	Fall	road_br	road	sand	scrub_ba	scrub_lu	shingle	sh_mo	silt	soft_cls	stream	he/ru	tot_res	tot_wood	urban	weir	max_wid	min_wid	wild
gr/he																							
altitude																							
am_mo																							
arable																							
artific																							
bed																							
bould																							
bridge																							
cobbles																							
conif_ba																							
conif_le																							
decid_ba																							
decid_lu																							
max_dep																							
min_dep																							
derelict																							
disturb																							
farming																							
grazed																							
unman																							
grass																							
hard_cls																							
Herac																							
human																							
impat																							
island																							
max_ht																							
max_slop																							
min_ht																							
min_slop																							
mixed																							
mud	0.134	-0.179																					
reeds	0.182	-0.213	0.297																				
Fall																							
road_br	0.177		0.228	0.165																			
road	-0.073		-0.103		0.113	0.309																	
sand	-0.095	0.072																					
scrub_ba	0.072	-0.168	0.213	0.134		0.068	-0.106																
scrub_lu			-0.099	-0.090		-0.089			0.152														
shingle	-0.098		-0.231			-0.085		0.248															
sh_mo	-0.144							0.075	-0.091		0.127												
silt		-0.202	0.251	0.329		0.112		-0.273	0.124			0.092											
soft_cls	-0.096		-0.098			-0.166	-0.077	0.206	0.081	0.103	0.296	0.138											
stream			-0.155						-0.108														
he/ru		-0.205	0.349	0.298	0.096	0.072	-0.083	0.077	0.220	-0.070	0.093		0.340		-0.148								
tot_res	-0.081				0.118	0.175	0.192	0.111				0.105			-0.067								
tot_wood	-0.113	0.260	-0.281	-0.263		-0.159		0.230		0.149	0.243		-0.188	0.206		-0.134	0.149						
urban	-0.078			-0.085	0.082	0.197	0.223	0.081				0.090			-0.078		0.797	0.139					
weir		0.096		-0.077		0.141	0.130		-0.096	0.070									0.105				
max_wid	-0.294	0.075	-0.378	-0.222	0.100	-0.313		0.111	-0.130	0.090	0.213	0.185	-0.169	0.213	0.082		0.150	0.364	0.185				
min_wid	-0.264	0.083	-0.378	-0.251	0.111	-0.322		0.120	-0.117	0.085	0.182	0.184	-0.158	0.228	0.085	-0.090	0.162	0.369	0.186		0.945		
wild		0.074	-0.282	-0.078		-0.115	0.071		-0.128	0.238	0.119	0.100	-0.087	0.153	0.140	-0.252		0.112		0.073	0.180	0.161	

Table AII.6 (cont.) Spearman rank correlation analyses between all RCS variables extracted from the Tees catchment RCS.